DOCUMENT RESUME

ED 394 958 SP 036 662

AUTHOR	Kansanen, Pertti, Ed.
TITLE	Discussion on Some Educational Issues VI. Research Report 145.
INSTITUTION	Helsinki Univ., (Finland). Dept. of Teacher Education.
REPORT NO	ISBN-951-45-7034-0; ISSN-0359-4203
PUB DATE	95
NOTE	212p.
AVAILABLE FROM	Department of Teacher Education, P.O. Box 38 (Ratakatu 6 A), 00014 University of Helsinki, Helsinki, Finland.
PUB TYPE	Collected Works - General (020) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120)
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	*Educational Change; Educational Legislation; Educational Practices; *Educational Research; Elementary Secondary Education; Foreign Countries; Higher Education; *Preservice Teacher Education; Program Evaluation; Reflective Teaching; Student Motivation; Student Teaching; *Teacher Education Curriculum; Teacher Evaluation; *Theory Practice Relationship
IDENTIFIERS	Denmark; Finland; Germany; Norway; United Kingdom

ABSTRACT

This research report presents 10 papers offering views of Finnish, Swedish, British, Danish, and United States educators on teaching and teacher education. The essays are as follows: (1) "The Missing Content in Teaching: Focus on What Teachers Reflect Upon While They Are Teaching" (Mikael Alexandersson); (2) "What Is Reflection? On Reflection in the Teaching Profession and Teacher Education" (Jan Bengtsson); (3) "Curiosity, Interest, and Intrinsic Motivation: A Conceptual Analysis" (Reijo Byman); (4) "The Development of Initial Teacher Education: Insights from Research on Learning to Teach" (James Calderhead); (5) "The Role of School Practice in Teacher Education" (Juhani Hytonen); (6) "Educational Knowledge and Reality" (Jujani Jussila); (7) "The 'Deutsche Didaktik' and the American Research on Teaching" (Pertti Kansanen); (8) "Concepts of Freedom in Danish School Legislation" (Sven Erik Nordenbo); (9) "The Denial of Change in the Process of Change: Systems of Ideas and the Construction of National Evaluations" (Thomas S. Popkewitz); and (10) "Evaluation of a Prototype Teacher Enhancement Program on Science Performance Assessment" (Maria Araceli Ruiz-Primo, Richard J. Shavelson, and Gail P. Baxter). (ND)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

h Repor

145

DISCUSSIONS ON SOME ECUCATIONAL ISSUES VI

Edited by Pertti Kansanen

Department of Teacher Education
University of Helsinki

ED 394 958

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
E. Niekanen
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

SP036662

BEST COPY AVAILABLE

Editorial Board:
Pirkko Anttila
Terttu Gröhn
Jarkko Hautamäki
Pirkko Anttila
Terttu Gröhn
Jarkko Hautamäki
Juhani Hytönen
Pertti Kansanen
Irina Koskinen
Veijo Meisalo
Kari Perenius (Secretary)

Available from:
Department of Teacher Education
P.O. Box 38 (Ratakatu 6 A)
00014 UNIVERSITY OF HELSINKI
Phone +358 0 191 8112
Fax +358 0 191 8114

Research Report 145

DISCUSSIONS ON SOME EDUCATIONAL ISSUES VI

Edited by Pertti Kansanen

Helsinki 1995

ISBN 951-45-7034-0
ISSN 0359-4203

Yliopistopaino
1995

Contents

Contributors 4

The Missing Content in Teaching
Focus on what teachers reflect upon while they are teaching
Mikael Alexandersson 5

What is Reflection?
On reflection in the teaching profession and teacher education
Jan Bengtsson 23

Curiosity, Interest, and Intrinsic Motivation: A Conceptual Analysis
Reijo Byman 35

The Development of Initial Teacher Education: Insights from Research on Learning to Teach
James Calderhead 63

The Role of School Practice in Teacher Education
Juhani Hytönen 77

Educational Knowledge and Reality
Juhani Jussila 85

The *Deutsche Didaktik* and the American Research on Teaching
Pertti Kansanen 97

Concepts of Freedom in Danish School Legislation
Sven Erik Nordenbo 119

The Denial of Change in the Process of Change: Systems of Ideas and the Construction of National Evaluations
Thomas S. Popkewitz 129

Evaluation of a Prototype Teacher Enhancement Program on Science Performance Assessment
Maria Araceli Ruiz-Primo, Richard J. Shavelson, and Gail P. Baxter 173

Contributors

Alexandersson, Mikael	Ph.D., University of Gothenburg, Sweden
Baxter, Gail P.	Assistant Professor, University of Michigan, USA
Bengtsson, Jan	Ph.D., Docent, University of Gothenburg, Sweden
Byman, Reijo	M.A., University of Helsinki, Finland
Calderhead, James	Professor, University of Bath, Great Britain
Hytönen, Juhani	Professor, University of Helsinki, Finland
Jussila, Juhani	Professor, University of Lapland, Finland
Kansanen, Pertti	Professor, University of Helsinki, Finland
Nordenbo, Sven Erik	Professor, University of Copenhagen, Denmark
Popkewitz, Tom	Professor, University of Wisconsin, USA
Ruiz-Primo, Maria Araceli	Research Associate, Stanford University, USA
Shavelson, Richard J.	Professor, Stanford University, USA

The Missing Content in Teaching

Focus on what teachers reflect upon while they are teaching

Mikael Alexandersson

This article is based on a study that has been developed within the research field (*Research on Teacher Thinking*) which focuses on how teachers reflect and think upon different aspects in their daily work.[1] The main aim for the study is to describe what teachers experience in their teaching and the central research question is " What do teachers direct their consciousness towards during their teaching?". A related purpose is to illuminate teacher's awareness in relation to their conception of their own working methods.

A phenomenological framework

One way to focus the central research question above is to take the point of departure in phenomenology. Phenomenology investigates that which appears to someone. This investigation can take place through studies of *how* people experience objects, phenomena or happenings in the world and *what* the experienced conceived to be. When something is experienced it has a *meaning*. This something appears in our consciousness (when we experience something), and the appearance occurs in a world which is historical, cultural and social. At the same time, however, we are historical, cultural and social beings. As Merleau-Ponty (1965) expresses it, a meeting between the individual and the world takes place with every act of consciousness. That meeting is the very condition of our existence. When we experience something in our existence, we do not experience isolated properties of an object, but we experience these properties – both functional and valuational – simultaneously, and we interpret their significance. My consciousness structures that which I experience but *what* my consciousness constructs depends upon my previous experiences. Experience consists of all the special thoughts, memories, emotions, expectations, etc., which the experienced elicits in us. As far as the present study is concerned, when a teacher experiences teaching it is an expression of an *interaction* between the teacher and the teaching itself.

---

[1] The article is partly a summary of a dissertation report (Alexandersson 1994a). A similarly version is published in Carlgren, Handal & Vaage 1994 (Alexandersson 1994b). In this article the presentation does not follow the same structure as in the dissertation report. For instance, the case studies are not presented here.

The theoretical foundation for the study is related to the phenomenological concepts of "reflection" and "intentionality". Reflection is here being used in terms of phenomenological reflection: The purpose of reflection is to try to grasp the essential meaning of something. According to Van Manen (1990, p. 77) "the insight into the essence of a phenomenon involves a process of reflectively appropriating, of clarifying, and of making explicit the structure of meaning of the lived experience". One of the fundamental questions in phenomenology concerns the implications of phenomena, which appear in consciousness, taking on a specific content. A teacher's experience of his or her own teaching, including, for example, perceiving, seeing, hearing or thinking about an object, corresponds to his or her intentionality and to the experienced object – *the intentional object*. The phenomenological concept – intentionality – could be defined to be *the directedness of an act to an object*. According to Idhe (1986) intentionality is the correlation between the mode of consciousness and the object.

To describe how teachers experience their own teaching

This study is carried out within the framework of a qualitative research approach called phenomenography (Marton, 1981, 1988; Alexandersson 1994c). Phenomenography attempts to describe and understand how people conceive, experience, perceive or understand different aspects of the world. This is done by investigating people's conceptions. From a phenomenographic point of view a conception is a way of seeing something. Phenomenography is based on the assumption that the conceptions of a single object differ among people. Differences in conceptions are explained by the fact that different people have different experiences due to their different relations to the world. People then make different analyses and arrive at different knowledge about the object concerned. Phenomenographic research attempts to describe these differences. There are three strategically important ways in which phenomenography is framed. In the content of the present study, these are as follows: there is a search for understanding variation in what teachers direct themselves towards in their own practice; instead of applying a model of description defined in advance, there is an attempt to explore the meaning of the variation in directedness; and there is an aim to find a meaning of the variation in directedness, by studying the teachers' experience of teaching.

The selection process of the teachers and the data collection

In qualitative studies, a general theoretical research problem is to find an acceptable point of balance between the requirement of a great range of variation and the requirement of a manageable empirical foundation. If the range of variation is too broad, the empirical foundation will be too confusing and difficult to handle. In the present study, the choice of teachers was made with the goal of achieving a balance between these two demands. The range of variation, for instance, includes a distribution relating to number of years of teaching experience, to a variation in the aim of the teaching, and also to the fact that the school context in itself should optimally vary in the teacher study. When considering generalisation in connection with the phenomenographical approach, this means that the conceptions which are attained shall cover as many qualities within the phenomenon as possible.

In order to optimize the possibilities of varying the methodology in the study, the choice of thirty teachers was made from the primary school[2]. The teachers were well recommended and respected for their teaching ability at their own school. This criterion included a range in different working methods. Technically a teacher in primary school can vary his, or her, method within most subjects. Most subjects in primary school can also comprise different proficiency factors, leading to a variation as regards verbal and visual performances. The educational material at this stage has a kind of layout that makes the methods vary considerably. Further, the variation of content in the teaching should not be constant in the study. With a variation also here, the following analysis could be made on a more general level. The selection process for this investigation can be described as "purposeful sampling" (Patton, 1990, p. 169) and as "purposive sampling" (Cohen & Manion, 1986, p. 100), together with what is known in ethnography as theoretical sampling. According to these principles for selection, the researcher hand-picks a group of subjects, using set criteria which correspond to the researcher's need for specific information.

In this study, data collection was limited to one occasion per teacher, but information was gathered on a continuous basis, with no predetermined limit to the number of teachers who would participate in the study. The data was processed on a continuous basis, which meant that I gradually became saturated with information. When my on-going analysis no longer disclosed new dimensions in the material, I therefore discontinued the collection of data. This occurred after twelve occasions. Glaser and Strauss (1967, p. 61 and p. 111) term this effect of data collection on the investigator as "theoretical saturation". The collection of data took an average of half a day for each teacher, and each session comprised three stages: video-recording, limited to one lesson, which varied between forty and sixty minutes; commentary, in which the teacher commented freely upon the documentary film (This stage constitutes the main source of information for this study.); follow-up discussion, a semi-structured interview for following up and examining more closely issues communicated in the commentary stage and for formulating general questions about the classroom teaching.

---

[2] In USA: Middle elementary school, in UK: Middle school or late primary school years, and in Sweden: mellanstadiet.

The purpose of video-recording the lesson was to stimulate the teacher's thinking about the completed lesson, to remind the teacher of his or her own thinking (the stimulated recall method, see Calderhead 1981; Peterson 1982). The question this study addresses – Towards what do teachers direct their consciousness when they teach? – indicates that the interest lies not in the video-recording as such, but in teacher thinking about the instruction which has been carried out. In other words, the focus is on what the teacher thought about during his or her teaching. The teacher was therefore requested to respond to the following three questions:

- What were you thinking about during the sequences you are now viewing?
- What were you doing?
- Why did you do what you did in that particular way?

Because the teachers could control the video-recorder by remote control during the commentary part, they could choose the sequences they wished to comment upon (389 sequences all together). Thus, they were able to stop the film to give spontaneous reactions to what they saw. That which was communicated through commentary is an intentional expression of what the teachers experienced in their own practice. In their descriptions of different occurrences, an intentionality lies in their way of giving shape to "that which appears" to them. That which they experience does not exist in explicit form but must be interpreted at a later stage on the basis of the totality to which their descriptions belong. The unstructured part was followed by a semi-structured interview. The purpose was to deepen and problematize the unstructured part. Teachers were asked the following questions:

- When commenting on certain aspects, scenes, incidents or sequences, why did you choose just those?
- Were you aware of your own acting and what was happening?
- Please describe your intention and in what way this documented hour was a part of a whole?
- How would you define your own working methods and of what importance are they to you?

Triangulation for in-depth analysis

For an in-depth study of how teachers experience their own teaching, triangulation was used both upon data collection – video-recording, commentary and follow-up discussion – and in connection with data analysis. In the field of behavioural science, triangulation,[3] which is usually described as a methodological approach utilising two or more methods (multi-method approach), aims at describing or explaining complex relations in a more fundamental way. The three methods of analysis used in this study are grounded in phenomenography, linguistics and qualitative analysis.

1. The Phenomenographic Analysis

A preliminary interpretation of how the teachers experience their own teaching was made simultaneous with data collection. To some extent, this interpretation directed the follow-up discussion which took place shortly afterwards. The interpretation was also an unavoidable part of the processing of the transcriptions of the commentary and follow-up and of the review of the video-recording. The analysis and interpretation of the commentary and follow-up discussion was divided into four phases, each having a different objective: to familiarize myself with the data and gain an overall impression; to note similarities and differences in the statements; to determine descriptive categories for conceptions; and to examine the underlying structure of the system of categorization.

---

[3] Literally, triangulation refers to a method used in the fields of navigation and military strategy. Using different positional points in nature, a particular point or position can be localised. See Patton 1990.

2. The Linguistic Analysis

In order to ascertain the manner and the extent to which the teachers commented upon specific phenomena and events, a large number of quotations from the teachers' statements were linguistically and statistically processed (a total of 581 quotations). The statistical analysis has treated the content of the running texts; this method is known as content analysis.[4]

3. The Qualitative Content Analysis

The aim of this content analysis was to develop an overview of the content of the direction of consciousness. To determine the content of these objects, all teacher comments were analysed. Recurring (frequent) topics and topics which the teachers regarded as important (significant) were noted (a total of 584 topics). By expressing the content of these topics in short phrases, complex reasoning could be simplified for the different cases. In this way, the analysis provided an illustrative and manageable overview of the content of all topics.

Results: Qualitative Variations in Directedness

The central finding in this study is the identification of the *direction-related* dimension. When the teachers directed their consciousness towards different aspects in the recorded teaching sequences, major qualitative variations could be identified as shown in Table 1.

Table 1. Qualitative variations in the direction-related dimension. Three qualitative main categories

A. Consciousness is directed towards **the activity itself**
B. Consciousness is directed towards **aims of general character**
C. Consciousness is directed also towards **a specific content**

The twelve teachers showed all three variations of directedness, but the main feature in each commentary could be located in one of them. A statistical analysis of quotations – 581 all together – was carried out in relation to these three categories of description. The contents of the quotations were examined both linguistically and for content, which resulted in a quantitative description of variations as regards the direction of consciousness. From the statistical point of view the quotations encompass three main component parts. These component parts were clearly separable; there was an activity and there was an object acted upon. The object could be stated in more general terms as aims of indefinite character and in more specific terms as a fixed and limited content. The findings indicate that, when commenting one specific instance, it was not natural that the teachers described the wholeness, i. e., their directedness in general was not towards a specific content in relation to a general aim and activity.

---

[4] Content analysis is commonly defined as a research method used to examine a broad range of issues in which a communicated content constitutes the very basis of analysis. The quotations or categories that exemplify an underlying idea, issue, or conception, for example, can be sought through the analysis. See, e.g., Holsti 1968.

MAIN CATEGORY A Consciousness is directed towards the activity itself (seven teachers)

What the different conceptions in this category had in common was the fact that consciousness was firstly directed to various aspects and situations which took place in the activity itself. There was no evident direction towards any general or concrete aim when the different instances were commented. Reflections consequently did not go beyond the visible activity. Apart from the topic of the content, it was the actual situation – the ongoing activity – which was commented upon. One can mention two levels of comments with the qualitative content, one level where only the activities in the different instances were described – the teachers referred to what was happening – and one level where the teachers explained what was happening. At the latter level the explanations were completed by statements about casual connections or about previous motives for their own or the pupils' activities. The analysis of the teachers' statements resulted in various conceptions forming a foundation for seven sub-categories.

Table 2 Sub-categories in main category A

A1. How the pupils are developed socially
A2. How a deep communication and relation is growing
A3. How the pupils are being noticed
A4. How to teach pupils to listen
A5. How a systematic teaching leads to activity
A6. How structured and balanced teaching is performed
A7. How do I think and how does the pupil think

One recurrent theme was the question of how a deep, social relationship with the pupils might be developed. When this theme was in focus, general or distinct aims were rarely drawn from the activities which were then noticed. In order to explain causes behind different sequences, the teachers based their explanations on experiences which had reference to an outside classroom situation – for instance to the pupils' social home situation. In several subcategories consciousness was directed towards situations where the teacher was helping pupils to discover knowledge themselves. When those situations were commented, statements often dealt with the importance of the fact that the individual pupil should have the possibility to develop his or her own thinking. In the following statement we can recognize the characteristics when the directedness is towards the ongoing process, i. e. how the teacher reflects upon her own thinking as well as upon the pupils' thinking.

> – I am trying to find out how they think here and how I am thinking. So I got some seconds there and took the opportunity. Something like that I was thinking.

Another striking feature was the fact that the consciousness was directed towards the teacher's own way of structuring and organizing the lesson. On this occasion the importance of systematic planning in order to make the lesson successful was emphasized. Through a stable and consistent structure on the lesson it would be possible to more effectively control the pupils' activities. Connected to the teacher's own part was also the effort to establish deep communication and a relationship with the pupils. Irrespective of the fact that the relationship was of social or of intellectual character, it could, according to the teachers, lead to an increase in the pupils' general understanding and learning.

The teachers in this main category had difficulties in commenting on the meaning of method in a distinct way. They preferred to refer to the possibility that the method should result through the development of co-operation between teacher and pupils during the lesson. As the method was more important than the content, the teacher's instruction could be separated from what was instructed about; i. e., what the pupils learned regarding content was not the main point in the instruction. They should rather learn methods which provided them with the independence they need when seeking knowledge. Methods in teaching were mentioned as a way to establish contact in terms of a good relationship with the pupils. One could trace the teacher's intention as being to capture of the attention of the pupils. In this perspective, the method became analogous to activities which satisfy the teacher's own needs and interests. Irrespective of arranged methods of any kind, the ultimate aim with the methods was to catch the pupils' attention.

In its concrete content, the method could be compared to areas like reading-methods and writing-methods, or to activities like developing pupils' thinking and ability to make reflections. When cognitive activities were commented, it was the variations in reflection and not the content in itself which were focused. Method was stated, partly as a means to organize the lesson so that the pupils should learn by being active themselves, and partly so that the pupils by their own discoveries might develop their thinking. In this way, the method was meant as a support to the pupils' cognitive development.

MAIN CATEGORY B. Consciousness is directed towards aims of general character (four teachers)

Teachers within this main category have a general attitude to the aims of different activities. In their directedness a forward leading aim can be found which goes beyond the activity itself, but does not aim at a specific outcome of the activity. The comments mostly deal with the fact that the actual activity has an aim of general character. Nearly half of the total amount of quotations can be assigned to descriptions of general aims. However, in the quotations there is no direction towards any specific content. The descriptions of aims are expressed at a more general level. The teachers have given reports of the situations rather than analyses. The following four qualitatively different conceptions can be distinguished concerning the content in direction of consciousness.

Table 3. Sub-categories in main category B

B1. Aims for the present conversation
B2. Aims for the open attitude in the teaching
B3. Aims for the teacher's active discipline
B4. Aims for catching the pupil's attention

The content in the descriptions of general aims concerned the teacher's ambition to develop relationships, to create high activity or to activate the pupils' thinking. The various methods which, directly or indirectly, were explicated in the commentaries and correlated to the clarification of method conceptions, made in the semi-structured interview, were really directed towards these descriptions of general aims. The recorded teaching sequence brought, throughout, thoughts on different principles for raising pupil activity. The directedness was towards visible, action-related situations which came from both the teacher as well as from the pupils. This started comments about the importance of action for education and knowledge in general. Reflection concerning the pupils learning was on very few occasions related to a fixed content but rather more to thinking as an activity. Thinking was, in this perspective, one of several activities in the teaching process. The statement below demonstrates the characteristics of the directedness in the category.

> – The aim is here to bring thoughts and questions with them, so that they all the time feel forced to activate a thinking process. That is the purpose of this question.

In the same way as in the previous main category, teachers held a more general attitude to the conception of method. They conveyed a certain irresolution in clarifying the method conception, and they connected the conception, at first hand, to their own activities and not to the pupils' learning. However, there was one divergence. The method conception seemed to be more reflected upon by teachers, in this main category.

MAIN CATEGORY C Consciousness is directed also towards a specific content (one teacher)

This third direction-related main category differs from the other two by the fact that consciousness is directed not only towards the activity in itself or towards a general aim, but also to a specific content. The direction of the quotations towards a general aim and towards a specific content regarding the present situation comprises two-thirds of the total number of quotations from the twelve teachers. A content-related discussion is principally always preceded by focusing the activity and/or the aims of general character. The present activity – alternatively the aim for this – is connected to content related intentions for the teachers' own activities as well as for the acts of the pupils. Of the twelve teachers only one represented this direction-related main category. The outstanding feature in the teacher's direction of consciousness corresponds essentially to the category "to activate the pupils' thinking". The qualitative conception is contained in the following sub-category:

Table 4. Sub-category in main category C

C1. That the pupils learn a specific content

The teachers' comments concerned mainly the reflection activity of the individual pupil and his or her ability to verbalize the reflection or to describe the reflection in a concrete action. Consciousness was then directed towards the way in which the pupil expressed his or her action in relation to a fixed content. There was an evident direction in the teachers' comment to produce a synthesis of, on the one hand, the pupils' reflection and action and, on the other, the teaching content. Reflection, action and content were consequently integrated. They were mutually dependent on each other and on each other's conditions. The three aspects in question were also expected to be correct. Neither in the unstructured interview nor in the follow-up discussion was there any statement which pointed out that a pupil's verbalized thought – in spite of quality – should be accepted and passed by the teacher. When the three aspects formed a correct totality, they were accepted. From this point of view the development of the pupils' thinking processes will become a part of a large totality. The following statement shows this reasoning concerning the activating of the pupils' thinking. The characteristics for this statement are the obvious directedness towards the pupils' learning of a specific content.

> – I'm walking to them all to see if they all know what a triangle is. Then some were a bit uncertain of it. They managed to make a loose angle, because then the whole of it is like an angle for him, but later when it is fitted in the triangle... If they are a bit away on the wrong track. Then I choose to walk between all of them and ask them to show me the angle.

If there was a connection between thought and action of the pupil, according to the teacher, that could mean an increase in the possibilities for the pupil to "think right". The "inner method", i.e., the pupil's cognitive activity, was comprehended as internalized in the pupil's consciousness. The "outer method", in terms of way of directions, became then means to facilitate and support such cognitive activities which could mean that the pupil's thought and action would be integrated into a totality. In this way, the method technically took a more remote place in the teacher's own consciousness.

Summary and discussion

In the table below, the results of the qualitative analysis and the statistical analysis of the contents of the quotations in each qualitative category are presented. In the horizontal heading the three statistical categories represent what the teachers made comments about. In category 1 the quotations are about activities – the present is described and explained. In category 2 the quotations are also about general aims for the activities and in category 3 the quotations include specific contents in relation to the aims and the activities. The statistical analysis revealed that the teachers in each qualitative category emphazised different component parts. For instance, the seven teachers which comprise main category A have their main feature in the statistical category 1. Out of a total of 347 quotations only 14 (4 percent) were in category 3. On the other hand, 266 quotations (78 percent) were located to the statistical category 1.

Table 5. Number of quotations in each qualitative main category

Qualitative Main categories		Statistical Categories 1	2	3	No. of quotations
A.	Consciousness is directed towards the activity itself (7 teachers)	266 (78)	67 (18)	14 (4)	347
B.	Consciousness is directed towards aims of general character (4 teachers)	87 (51)	52 (30)	32 (19)	171
C.	Consciousness is directed also towards a specific content (1 teacher)	23 (37)	12 (19)	28 (44)	63
	Total no. of quotations	376 (65)	131 (22)	74 (13)	581

Note: (Per cent)

The overview above indicates that, of a total of 581 quotations, 376 (65 percent) concerned the activity – either the ongoing process or the function of the teacher's own activity in the classroom. In 131 quotations (22 percent) the directedness concerned aims of general character and in 74 quotations (13 percent) a specific content.

An issue of central importance is how the three main categories stand in relation to the group of teachers and the subject taught, which formed the object of the analysis. Regardless of school, grade, subject of instruction and content of subject, the teachers commented upon similar aspects of their practice – above all, capturing their pupils' attention and developing the pupils' thinking or creating a sense of community in the class. However, although the teachers direct their consciousness towards similar aspects, they do so in qualitative different ways. Thus, the subject or subject matter does not determine the directedness of the teachers in this study. Data was collected upon a single occasion for each teacher and from a specific research perspective. If a different classroom session had been documented, different results would have emerged for each individual teacher. I can confidently state that the three main categories – and possibly others – exist irrespective of the concrete teaching context. These categories appear upon different occasions, and teachers – regardless of the grade or subject they are teaching – can then be carriers of the structural content in each category.

It is not possible to argue that these differences in directedness are general for teachers in primary school, but on the contrary the results show that this kind of different directedness among teachers in their own teaching process may take place. In fact, how the teachers reflected upon their own practice points out the fact of obvious variations in directedness. One could say, that the overarching theme in this study indicates that highly skilled teachers not to be necessarily driven by the aim that their pupils should develop certain specific understanding, knowledge or skill; the teachers were in general not directed towards some specific content of the pupils learning. The study also points out differences in the way in which the teachers comprehended their own working methods. With regard to this study, it seems to be a correspondence between how consciousness was directed and the conception of working methods. When directedness pointed towards some specific content for the pupils' learning, the method technically took a more remote place in the teacher's consciousness. When the activity itself was in focus, the method was more important than the content. The outcome in terms of specific understanding, knowledge or skill was not the main point when different methods were used in the teaching process.

The general picture is quite clear and could be summarized in the following way: A specific content does not appear to be a major driving force for the teachers' activity. The teachers in general were not oriented towards specific learning aims and they rarely focused on means-ends relations.

Comments on the results

On the result basis three central themes will here be commented; different conceptions of method, different conceptions of thinking and different conceptions of content.

1. Different conceptions of method

The majority of teachers proved to have obvious difficulties in abstracting the concept of method from their concrete teaching practice. Whether or not the method is planned, its purpose is often to engage the pupils' attention and to establish a stable teaching situation, which primarily depends on the structure of the lesson. The method constitutes the means for the teacher to realise his or her intentions and to meet his or her need for a high level of activity and concentration among the pupils. This conception corresponds to an *outer method*. On the other hand, attention was drawn to those methods which the pupils themselves need to develop in order to benefit from instruction; these are *inner methods*. Methods of instruction aim at developing skills such as reading and writing or thinking. Methods are then said to be the means of organising instruction so that pupils can learn through their own activities and can develop cognitively through their own discoveries.

2. Different conceptions of thinking

A recurrent theme expressed by teachers concerned the thought processes of pupils. When consciousness was directed towards pupil thinking, two different conceptions of this could be identified. When intentions such as getting students to think were expressed, teacher activity was geared towards capturing the students' attention to activate their thought processes: If pupils thought, the teaching goal was met. The thought processes of pupils were considered capable of being developed through, for example, systematic and structured instruction. In these cases, the teachers' point of departure did not lie in the pupils' understanding of a specific content; rather, their own understanding of the content was in focus. Implicit to the latter conception, which was found in the third main category, the idea is that thinking is a part of a whole. The qualitative difference between ways of discussing the act of thinking lies in the conception of relationships between the content, the general aim and the activity in progress. When the act of consciousness is directed towards pupil thinking, thinking is conceived of as means of bringing together the "right" thought with the "right" action. When teachers reflected upon their own thinking, however, the pupils' comprehension was always focused using a specific content.

3. Different conceptions of content

Between the individual descriptive categories of the third main category and the individual descriptive categories of the other two categories there is a systematic difference in what *content* is considered to be. This difference did not appear in the beginning of the analysis but gradually emerged from the material. When the consciousness of the teachers was directed towards a specific content, two separate aspects of that content became apparent. On the one hand, content can correspond to x, y or z which the pupils must learn. On the other hand, content can refer to a *thought* content; the point of departure is then the pupils' thinking about a specific content. The systematic difference lies precisely in the fact that the teacher who carries the structured content in the third main category assumes that his or her own thought and that of the pupils comes into *contact* when a specific content is being communicated.

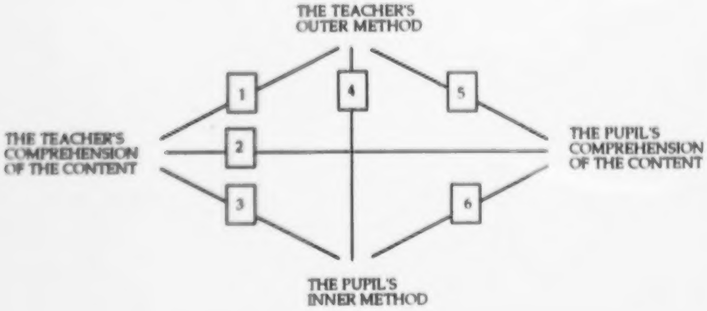
Towards a dialectical model of the relationships between content and method

The findings of this study indicates that the specific content of teaching must be discussed from different perspectives. This is true for, among other things, discussion of questions about how a specific content should be structured and organised on the basis of student comprehension of the content and about how student comprehension can be forwarded. Here, recent research on how students think about a specific content is useful. Viewing instruction from the pupils' perspective entails analysing how they deal with a specific content in relation to a set teaching goal. The *inner* method – the method found within the pupils – then corresponds to how pupils understand and treat a specific content. The way pupils define their own actions in a formal learning situation should thus be central to both the *inner* and *outer* method. It is the students' understanding of a specific content which constitutes the basis of the construction of method at both levels.

I believe that, to enhance teaching and learning, the teacher must not only seek knowledge about the pupil but must also seek the knowledge the pupil already has. This collected knowledge can then serve as a basis for the teacher's activities. The way in which the pupil understands and handles a specific content may become the teacher's method. There is thus an interaction between the teacher's activities and the pupil's method, for these exert a mutual influence which can be expressed as a dialectical relationship between outer and inner methods.

Assuming, however, that the logic of this method is inextricably bound to a particular situation, the method may require situating. The method arises as an interaction between a specific content, the pupil's comprehension of the content and the teacher's awareness of this interaction. Based on this assumption, the method of instruction can be grounded in a dialectical model made up, on the one hand, of the teacher's knowledge of the content and the pupil's comprehension of the specific content and, on the other, of the outer and inner method, respectively. (In the figure below, the following six relationships appear.)

Figure 6. A dialectical model of the relationships between content and method


THE TEACHER'S OUTER METHOD
1
4
5
THE TEACHER'S COMPREHENSION OF THE CONTENT
2
THE PUPIL'S COMPREHENSION OF THE CONTENT
3
6
THE PUPIL'S INNER METHOD

The question is if teaching may, in fact, first be said to occur when a teacher has knowledge about a content and method in relation to the individual student. Such a definition places a demand for awareness of how the individual student understands a content, how the student can use his or her understanding of the content (the inner method) and what the teacher can do to help develop this process. The teaching method thus becomes adjusted to the teacher's understanding of the specific content and to the way the student understands it. At least four didactic requirements for knowledge are then generated: Knowledge about the specific content, knowledge about how the student thinks of this content, knowledge about how the student's comprehension of content can be supported and *awareness of the importance of this interactive process*.

References

ALEXANDERSSON, M. (1994a) *Metod och Medvetande (Method and Consciousness)*. Doctoral dissertation. Göteborg Studies In Educational Sciences 96. Göteborg: Acta Universitatis Gothoburgensis.

ALEXANDERSSON, M. (1994b). Focusing Teacher Consciusness. In G. Handal, I. Carlgren & S. Vaage (Eds.), *Teachers' Minds and Actions*. London: The Falmer Press.

ALEXANDERSSON, M. (1994c). Den fenomenografiska forskningsansatsens fokus. In B. Starrin & P - G. Svensson (red.), *Kvalitativ analys, vetenskap och samhälle*. Lund: Studentlitteratur.

CALDERHEAD, J. (1981) Stimulated Recall: A Method for Research on Teaching. *British Journal of Educational Psychology*, 51, 2, 211 - 217.

COHEN, L. & MANION, L. (1986). *Research methods in education*. London: Croom Helm.

GLASER, B. & STRAUSS, A. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine Publishing Company.

HOLSTI, O. R. (1968). Content Analysis. In G. Lindzey & E. Aronson (Eds.), *The Handbook of Social Psychology. Volume 2: Research Methods*. Reading, Mass.: Addison-Wesley.

IDHE, D. (1986) *Experimental phenomenology*. New York: Putnam.

MARTON, F. (1981) Phenomenography – describing conceptions of the world around us. *Instructional Science*, 10, 177 - 200.

MARTON, F. (1988) Phenomenography: Exploring Different Conceptions of Reality. In D. Fetterman (Ed.), *Qualitative Approaches to Evaluation in Education*. New York: Praeger.

MERLEAU-PONTY, M. (1965) *Phenomenology of Perception*. London: Routledge & Kegan Paul.

PATTON, M. Q. (1990) *Qualitative Evaluation and Research Methods*. Newbury Park, California: Sage Publications.

PETERSON, P. L. (1982) Beyond Time on Task: Students' Reports of Their Thought Processes during Classroom Instruction. *The Elementary School Journal*, 5, 481 - 491.

VAN MANEN, M. (1990) *Researching lived experience. Human science for an action sensitive pedagogy*. Ontario: The Althouse Press.

What is Reflection?
On reflection in the teaching profession and teacher education

Jan Bengtsson

I The notion of reflection in contemporary pedagogy

The actual situation in research and practice

During the last decade, the interest in reflection has grown massively in Anglo-Saxon pedagogy and has become a key-concept in discussions about teacher education and the teaching profession.

A possible explanation of the fact that this development has mainly occurred in Anglo-Saxon countries and not in German- and French-speaking countries on the Continent of Europe is the positivist orientation in the social and human sciences. This orientation, although not all excluding, has been dominant in the Anglo-Saxon countries, and it has a strongly instrumental view of the relation between science and professional practice. In the German- and French-speaking countries, the situation has not been as dominated by positivist-instrumentalistic traditions ("tradition" in plural because positivism is no more homogenous than other traditions). In these countries phenomenology, existentialism, hermeneutics, critical theory, structuralism and other traditions with mostly a quite different view on the relationship between science and professional practice have been at least as significant. For this reason it is not unexpected that much of the present Anglo-Saxon discussion on reflection is to a large extent supported by the latter traditions. During the last decade, American pragmatism, originally strongly influenced by Continental philosophy, has come into favour again in Anglo-Saxon pedagogy.

One important inspiration for the widely spread interest in the notion of reflection has been Donald Schön's book *The reflective practitioner* (1983), and this in spite of the paradoxical circumstance that Schön's book does no treat the teaching profession directly although he is professor in urban studies and education. In subsequent books, *Educating the reflective practitioner* (1987) and *The reflective turn* (1990), however, pedagogy has been considered. Other theoretical works of great importance for this development are Max van Manen's article "Linking ways of knowing with ways of being practical" (1977) and John Dewey's *How we think*, in its second and completely revised edition from 1933, which has been rediscovered and gained new importance.

Today, it is hard to keep up with the continuos stream of new contributions on reflection in the teaching profession. In pedagogical journals, new articles are steadily published, and special issues are devoted to these fields, e. g. *Theory into Practice* (1990/3) and *Journal of Teacher Education* (1989/2). During recent years a large number of anthologies and monographs have also been published. Some examples are: Grimmett and Erickson (eds.) *Reflection in teacher education* (1988), Waxman, Freiberg, Vaughn and Veil (eds.) *Images of reflection in teacher education* (1988), Clift, Houston and Pugach (eds.) *Encouraging reflective practice in education* (1990), Tabachnick and Zeichner (eds.) *Issues and practices in inquiry-oriented teacher education* (1991), Valli (ed.) *Reflective teacher education* (1992), Russell and Munby (eds.) *Teachers and teaching* (1992).

The interest has, however, not stopped at discourses on reflection. Different efforts have also been made to introduce reflection into teacher education. University of Wisconsin in Madison has in this connection played a leading role. Kenneth M. Zeichner, professor at that university, has in a serial of articles such as "Reflective teaching and field-based experience in teacher education" (1981), "Teaching student teachers to reflect" (1987) and "Preparing reflective teachers" (1987) described and discussed the content and outcomes from reflective-based education in Madison. He has also referred to an extended interest on reflection at other institutions in the USA and Canada (Zeichner and Tabachnick 1991, p. 1), as has Calderhead (1989, p. 43) regarding developments in England and in Australia.

Corresponding to the integration of reflection into teacher education, the publication of textbooks has started. Some examples are Pollard and Tann: *Reflective teaching in the primary school* (1987), Posner: *Field experience: A guide to reflective teaching* (1985), Cruickshank: *Reflective teaching* (1987), Henderson (ed.) *Reflective teaching: Becoming an inquiring educator* (1992) and Ross, Bondy and Kyle: *Reflective teaching for student empowerment* (1993).

BEST COPY AVAILABLE

The problem situation

Donald Schön seems to have found the proper expression for the present situation within teacher education and the teaching profession in a recently published anthology with the title *The reflective turn* (1990). Reflection is today on everybody's lips, and this has created the paradoxical situation that "reflection" is often used in an unreflected manner. This doesn't only apply to the public debate. Also in written contributions, a notion of reflection is used that seldom is clarified such that the term "reflection" is hiding a manifold of different ideas on the nature of reflection and its implications for the teaching profession and teacher education. This situation is partly mirrored in the jungle of expressions that circulate, and it catches at the same time something about the way reflection is understood as well as the hopes connected with its consequences. Some of the occurring expressions are "reflective teaching", "reflective practice", "the teacher as researcher", "the teacher as decisionmaker", "reflective thinking", "the teacher as problemsolver", "inquiry-oriented teacher education", "reflection-in-action" and "the teacher as professional".

Already from these expressions it is evident that reflection is understood in many different ways. (1) It is something that occurs in action. (2) It is separated from action and is of another kind, viz. a cognitive activity. (3) It is itself an action, but of another kind than teaching action, viz. a kind of self research.

It is also possible to see indications of what reflection is supposed to lead to. In all cases reflection is supposed to have a certain enlightenment function. It shall, to paraphrase Kant (1981, p. 53), declare the teachers maturity so that they can act independently. But at the same time there is a great variation concerning the degree of supposed independence. Least far-reaching is the thesis of making the teacher independent of science. According to this thought, the relationship between science and teacher is thought of as a relation between a passive teacher who applies scientific knowledge, and where reflection gives the teacher self-control over his or her profession. Most far-reaching is the idea of a completely autonomous teacher who, with the help of reflection, is able to see through all political, social, historical and other ideological factors embedded in every educational situation and from this elevated position chooses freely and consciously in order to take full responsibility for his or her actions.

In the same way as the answer of the question what reflection can lead to is dependent on the determination of reflection, there is also dependence between the conception of reflection, teacher competence and teacher education. If professional competence is supposed to be an absolute autonomy of the above described kind, it seems to be natural to give the student teacher a thorough training in critical reflection before he or she is permitted to start practical teacher training in school. Otherwise, the teacher will be the victim of the routines and the presuppositions of the professional practice. If, on the contrary, the professional competence is supposed to be the capacity to reflect in action, this can be used as an argument for first of all giving the teacher student a supervised practical training in order to initiate him or her into the practical problems of the profession.

The above described situation gives rise to a number of problems. First of all, the divergent uses of the term "reflection" indicate that it is fundamentally unclear what reflection really is. It is sometimes inevitable that the question presents itself if the authors are talking about the same thing. This ambiguity is consequently not of the good kind, to use the words of Merleau-Ponty (1960, p. 10), that allow the different meanings to complement each other. Reflection doesn't have all the properties it is attributed in the literature. Therefore, it is necessary to discern the properties it has from those it doesn't have.

All determinations of reflection have specific consequences. Independently, whether or not the determination is rendered explicit or left implicit, it indicates possible uses of reflection in the teaching profession and in teacher education. All possible peculiarities, exaggerations etc. which the notion of reflection may contain have consequences for the intended uses of reflection. Therefore, it is necessary to estimate the possible consequences of reflection.

My description of the present problem situation is not exhaustive, but it is my hope that I have succeeded in formulating the importance of a group of fundamental problems, all of which are of philosophical and not empirical character.

II Some basic determinations of reflection

The contribution of this part of the article is again limited. The intention is not to present an exhaustive determination of the nature of reflection, nor to solve all problems connected with the use of reflection in the teaching profession and teacher education. I limit my contribution to the illumination of some elementary features of reflection.

The etymology of the term "reflection"

It seems reasonable to start an inquiry concerning reflection with an etymological examination of the term "reflection". However, in this context the aim is not to re-introduce an overlooked meaning in the history of the term which now is going to be brought forth as the explanation of what reflection is. This kind of explanation can certainly have a value and in many cases can be justified if it demonstrates some unnoticed aspect of reflection. Language and world are, to be sure, interdependent, something everybody who speaks a foreign language learns as they discover that words, phrases, expressions etc. are not always possible to translate with retained meaning and sometimes have no equivalence at all.

An interesting example is the German word "Schüler" and the French "élève", both of which are translated into English as "pupil". Although, all of the words have the same reference, they have different content of meaning. They focus on different aspects of being a pupil. The German word focus on training. "Schüler" (pupil) is connected with "Schule" (school) and the verb "schulen" (train). To be a pupil in Germany is, accordingly, connected with being trained in a school situation. In French, on the contrary, the accent is on elevation. "Élève" (pupil) is connected with the verb "élever" (elevate), but has nothing to do with "école" (school). To be a pupil in France has, accordingly, the meaning of being elevated to constantly higher levels. By translation of the German and the French word to French and German respectively the meaning of the word is changed as well.

My motive for starting with an etymological inquiry is, then, not to uncover an overlooked meaning in the word "reflection". The motive is rather the opposite. The word "reflection" has a latin origin and is used broadly (French "réflexion", German "Reflektion", Swedish "reflektion") with a common meaning that doesn't seem to have changed much over time. This meaning has, moreover, the big advantage that the core meaning of the term corresponds with the core properties of reflection. It is, however, important to observe that it is a core meaning and not an exhaustive determination of reflection. It is this broad consensus concerning a pertinent choice of words that motivates an etymological starting point.

Reflection as self-reflection

The word "reflection" originates from the latin verb "reflectere" which means bend or turn ("flectere") backwards or back ("re"). The term was originally introduced in optics to describe the reflection of light against a smooth water surface, a mirror or such like. In English there doesn't seem to be a better way to express the return of light from water or a mirror than by the word "reflect".

This meaning is also kept when the word is used in human contexts. In this context, however, the word is seldom used in a literal sense, i. e. with the meaning of physical self-mirroring. Narcissus' seeing of his own mirror image in the water or the child's recognition of itself in the mirror are not examples of the meaning of the word "reflection" in human contexts. Certainly we can say that the face of Narcissus is reflected in the water, but this physical sense is not that in which the word "reflection" has obtained its widespread and established meaning. The reason for this is evident already in the case of Narcissus. A pure physical description of the relationship between Narcissus and his mirror image in the water is insufficient as well as uninteresting for catching the existing relationship. A description of this kind has not caught the new, transphysical and human in the relationship. The fact is that Narcissus recognizes himself, i. e. he doesn't see a beautiful image by which he is spellbound, but he is fascinated by recognizing himself as a person in the physical image on the surface of the water.

"Reflection", therefore, is mainly used metaphorically in human contexts. At least three reasons for this can be alleged. (1) Firstly, man is not passively reflected in the mirror. Man is him- or herself an active initiator and looks upon him- or herself in the mirror. (2) Secondly, there is, after all, no need for a mirror at all. Man turns him- or herself toward him- or herself and discovers him- or herself in the reflection without the help of something different and external. (3) Thirdly, that which is mirrored is not of the physical kind or, at least, not only physical. Instead, man discovers him- or herself as a person, his mental activities, but also the existential aspects of his life. A crucial question is, therefore, how subject and consciousness is understood, in our case that of the teacher and teacher student.

Reflection as thinking

The word reflection is also used with a different meaning in human contexts. With "reflection" is sometimes meant meditation, thinking, cogitation and similar intellectual activities. In this sense, "reflection" means that some phenomenon is subjected to thorough consideration, that thought dwells a longer period of time on an object in order to get a better and deeper understanding of it.

The object of reflection can be of two different kinds. It can be (1) one's own activity, but it can in principle be (2) any kind of object. Accordingly, in this sense it is possible to reflect upon something quite different than oneself, e. g. the origin of the earth, the distribution of the species, the nature of electricity, as well as upon one's own professional activities. One's own professional activity can be thematized by the reflective thought in different ways. The reflection can for instance aim at the planning of a lesson or at the evaluation of a carried out instruction, i. e. thinking can precede as well as succeed action.

But there are also types of objects between the other and the own. To this group belongs reflection upon (3) one's own profession, independently if it is teaching, research or some other professional activity that is aimed at. This field of reflection must not be identified with one's own professional activity, because one's own professional field is much vaster than one's own practice of the profession. Within the teaching profession thousands of other people have been working, in the past as well as in our time, in many different countries and cultures, and this is probably going to continue even in the foreseeable future. Reflection upon one's own professional field includes, consequently, many things which are different than one's own professional activity and which reflection upon one's own professional practice will not uncover.

Reflection upon one's own professional practice is limited to one's own individual practice whereas reflection upon the professional field includes super individual components such as the historical development of the profession, what it is at present, future possibilities etc. A crucial question in this connection is how the individual and the super individual should be understood and how they are related to each other.

At the same time as one's own professional field can not be identified with one's own professional practice it is evident that this field of reflection is not completely different from one's own professional practice in the same way as reflection upon the nature of electricity is. One's own professional practice is still a subset of one's own profession.

Reflection as self-understanding

The first and the second notions of reflection, reflection as self-reflection and reflection as thinking respectively, do not in many cases have very much in common. This applies in particular to the second notion of reflection in the second sense mentioned above, viz. reflection as thinking on an object without any backwards directed reference at all to the person who reflects. In this case, it is not a matter of real reflection, but it is simply thinking, meditation or the like.

But one must not be led astray by the fact that the two other types of object for reflection in the sense of thinking or meditation mentioned above, viz. one's own professional activity and one's own profession, have self-reference. Although thinking about one's own profession has an indirect reference to the reflecting subject, it doesn't attain selfdiscovery. In other words, it isn't a true reflection. But not even when one's own professional activity is made the object of thinking has thinking led to self-discovery. On the contrary, thinking presupposes that the true re-flection has already attained self-discovery.

If thinking is preceded by self-discovery, then the second notion of reflection can complement the first notion of reflection, because if the first kind of reflection leads to self-discovery, it seems to be a natural continuation to dwell upon the discovered in order to understand it and gain knowledge about it. In other words, the selfdiscovery is first accomplished by thinking on that which appears.

The distanciative function of self-reflection

As indicated above, reflection is never a normal state of man. Firstly, life is not only thinking; at best thinking is something constantly recurring. In the teaching profession, thinking is a recurring activity, but mostly, the teacher acts and perceives in the practice of his or her profession, and that is something different than thinking.

Secondly, life is even less directed against itself. The natural attitude of man is directed forwards against that which happens and exists in the world around him or her. This applies, of course, equally to teachers. They are busied toward all that happens around them in the classroom and by their teaching. Therefore, it is not paradoxical that that which is most close to us, viz. our selves, at the same time can be the most strange for us. The reason is that we have a natural distance to that which happens in the world. It is there in front of our nose. To ourselves, on the contrary, we don't have the same natural distance. But with the aid of reflection, the teacher can introduce distance to him- or herself and his or her activities, and thereby get sight of him- or herself.

By that is not said that the teacher by reflection can learn everything about himself and his professional activities or that reflection is the only way to get knowledge about oneself. Other possibilities are dialogue with colleagues and research (see Bengtsson 1993). The limits of selfreflection don't, however, belittle its importance as a knowledge instrument for the teacher in his or her professional development. But it is problematical to assert self-reflection as the teacher's only possibility for developing his or her professional practice, because thereby the limits of professional knowledge become identical with that of self-reflection. Instead, it ought to have the function as one of several possibilities of professional development.

An interesting problem in the connection with man's natural direction towards something different than him- or herself is how he or she can discover him- or herself at all. If he or she hadn't the slightest idea about him- or herself, there would be nothing that could lead him thereto. He or she would be imprisoned in the natural attitude toward the world. Man's way out of his or her immersion in the world and practice is a simultaneous awareness of him- or herself. In everything that we see and do, there is a certain consciousness of ourselves. The teacher that is completely absorbed by his or her teaching has nevertheless a marginal consciousness that it is he or she that teaches, i. e. a consciousness of oneself as a subject. This phenomenon is traditionally called "apperception", which literally means that the experience is accompanied by another experience or that something is co-experienced in the experience. And the co-experienced is in this case the experiencing subject. Thanks to apperception we can always choose to turn ourselves towards ourselves, i. e. to reflect.

Apperception is, consequently, not identical with self-consciousness. It is a presupposition for reflection and self-consciousness. Schön's influential but most unclear concept of reflection in action can in some passages of *The reflective practitioner* (1983) be understood as a confusion of apperception and self-consciousness. Apperception is in that case understood as a kind of reflection that continually accompanies the actions of the professional practitioner.

But this is only one of many possible interpretations that the book gives rise to. Sometimes "reflection in action" appears to be a confusion of reflection and reaction on and/or interaction with the situation which the practitioner is in. But to react on or interact with a situation is not the same as to reflect upon it – neither in the sense of "self-reflection" nor in that of "thinking".

In any case, it is evidently clear in Schön's account that with "reflection in action" he doesn't mean "reflection on action". This is explicitly and repeatedly stated. But how should, then, this reflection be thought? "Reflection", in both senses indicated above, i. e. "self-reflection" and "thinking", are reflections on something.

In some text passages the term "reflection in action" is still used in such a way that it must be understood as reflection on action. Thereby, we are back at reflection as a possibility for distanciation. But Schön's concept seems in this interpretation more readily to fit "reflection" in the sense of "thinking" than "self-reflection". This is also supported by the subtitle of the book: *How professionals think in action*. This heading could certainly give rise to a suspicion about an intellectual theory of action, i. e. a theory that makes action into an intellectual activity. There is, however, no support in the book for this interpretation.

When would it, then, be possible to make use of reflection on action? There are at least three natural opportunities where thinking can be related to the actions of the teacher. (1) The teacher can think before the action, e. g. planning a lesson. (2) The teacher can think after the action, e. g. evaluate a course. (3) The teacher can interrupt the action and think when problems arise.

In none of these three cases is it obvious that the teacher him- or herself and his or her activities are going to be thematized. A presupposition for this is that the thinking is preceded by self-reflection.

Acknowledgments

The present study was supported by grants from the Swedish Council for Studies of Higher Education and the National Agency for Education. I am also indebted to Dennis Beach at the Department of Education and Educational Research at the University of Göteborg for checking my English.

References

Bengtsson, Jan (1993): "Theory and practice. Two fundamental categories in the philosophy of teacher education", *Educational Review* (45/3), pp. 205-211.

Calderhead, James (1989): "Reflective teaching and teacher education", *Teaching and Teacher Education* (5/1), pp. 43-51.

Clift, Renée T., Houston, W. Robert and Pugach, Marleen C. (eds.) (1990): *Encouraging reflective practice in education. An analysis of issues and programs.* (New York and London: Teachers College Press).

Cruickshank, Donald R. (1987): Reflective teaching. *The preparation of students of teaching.* (Reston, Virginia: Association of teacher educators).

Dewey, John (1933): *How we think. A restatement of the relation of reflective thinking to the educative process.* (Boston: D. C. Heath and Comp.).

Grimmett, Peter P. and Erickson, Gaalen L. (eds.) (1988): *Reflection in teacher education.* (New York and London: Teachers College Press).

Henderson, James G. (ed.) (1992): *Reflective teaching: Becoming an inquiring educator.* (New York, Oxford etc.: Maxwell Macmillan).

Journal of Teacher Education 40 (1989/2). Critical reflection in teacher education. Practices and problems.

Kant, Immanuel (1981): "Beantwortung der Frage: Was ist Aufklärung?". In *Werkausgabe.* Edited by Wilhelm Weischedel. Volume XI. (Frankfurt am Main: Suhrkamp), pp. 53-61.

van Manen, Max (1977): "Linking ways of knowing with ways of being practical", *Curriculum Inquiry* (6/3), pp. 205228.

Merleau-Ponty, Maurice (1960): *Eloge de la philosophie.* (Paris: Gallimard).

Pollard, Andrew and Tann, Sarah (1987): *Reflective teaching in the primary school. A handbook for the classroom* (London: Cassell).

Posner, George J. (1985): *Field experience: A guide to reflective teaching.* (New York and London: Longman).

Ross, Dorene Doerre, Bondy, Elizabeth and Kyle, Diane Wells (1993): *Reflective teaching for student empowerment. Elementary curriculum and methods.* (New York: Macmillan).

Russell, Tom and Munby, Hugh (eds.) (1992): *Teachers and teaching. From classroom to reflection.* (London: The Falmer Press).

Schön, Donald A.(1983): *The reflective practitioner. How professionals think in action.* (New York: Basic Books).

Schön, Donald A. (1987): *Educating the reflective practitioner.* (San Francisco: Jossey-Bass).

Schön, Donald A. (ed.) (1990): *The reflective turn. Casestudies in and on educational practice.* (New York and London: Teachers College Press).

Theory into Practice 29 (1990/3). Teacher as researcher.

Valli, Linda (ed.) (1992): *Reflective teacher education. Cases and critiques.* (Albany, N. Y.: State University of New York Press).

Waxman, Hersholt C., Freiberg, J. C., Vaughn, J. and Veil, M. (eds.) (1988): *Images of reflection in teacher education*. (Reston, VA: Association of Teacher Educators).

Zeichner, Kenneth M. (1981): "Reflective teaching and field-based experience in teacher education", *Interchange* 12, pp. 1-22.

Zeichner, Kenneth M. (1987): "Preparing reflective teachers: An overview of instructional strategies which have beenemployed in preservice teacher education", *International Journal of Educational Research* (11/5), pp. 565-575.

Zeichner, Kenneth M. och Liston, Daniel P. (1987): "Teaching student teachers to reflect", *Harvard Educational Review* (57/1), p. 23-48.

Zeichner, Kenneth M. and Tabachnick, B. Robert (eds.) (1991): *Issues and practices in inquiry-oriented teacher education*. (New York and London: The Falmer Press).

Zeichner, Kenneth M. and Tabachnick, B. Robert (1991): "Reflections on reflective teaching", in Tabachnick, B. Robert and Zeichner, Kenneth M. (eds.): *Issues and practices in inquiryoriented teacher education*. (New York and London: The Falmer Press), pp. 1-21.

Curiosity, Interest, and Intrinsic Motivation: A Conceptual Analysis

Reijo Byman

Introduction

> Take a student, place him in a situation of moderate *uncertainty* [italics added] about some topic and get out of his way while he gets excited and attentive and directs his exploration to the source of his uncertainty. Moreover, research has demonstrated that he will enjoy his exploration and the accumulation of knowledge. (Day, 1982, p. 19)

This is Day's "recipe" to the classical problem "What can I do to motivate students to learn?". The central idea of this recipe is that uncertainty elicits curiosity and interest and thus starts a search for information. However, in practice the situation may not be so simple as Day suggested because an overdose of "uncertainty" can evoke anxiety instead of curiosity and interest. Thus, as Kansanen (1986) and Weiner (1992) have noted, "tricks" or "cookbook answers" are unrealistic solutions to the classical motivation problem. What is needed in motivational research according to Weiner (1990, 1992) is broad conceptual work which links the three concepts of motivation, cognition, and emotion.

Since Aristoteles, one of the ideal motivational teaching strategies has been to get students to explore, discover and learn actively. In recent years several psychologists and educators have preferred learning motivation which is free from the pressure of grades, extrinsic rewards and control. That is, they advocate motivation where learning activities are undertaken for their own sake. But what then makes a person gather information for its own sake - for the pleasure and satisfaction derived from the activity? One answer has been curiosity, another interest, and a third intrinsic motivation. On the other hand, schools are often blamed for fostering the wrong kind of motivation, for example, killing the natural curiosity of children. Is this because the word *curiosity* evokes both good and bad feelings in our minds or because teachers do not know what curiosity really is and for that reason cannot teach, as Day (1982, p. 20) expressed it, "the Curious Way". Is it so that the natural curiosity of children must be fostered not suppressed? Day and Maynes (1972) stated as follows:

> Each of us is born with a disposition to be curious, a disposition to know. Whether this curiosity survives and serves us effectively depends on the strategies used by our parents and teachers. These strategies are independent of the place of learning. (p. 69)

Curiosity, interest, and intrinsic motivation are supposed to have very similar positive effects on learning (cf. Bruner, 1966; Berlyne, 1965; Deci & Ryan, 1985; Dewey, 1913; Schiefele, Krapp & Wintler, 1992). Moreover, Iran-Nejad, McKeachie, and Berliner (1990) have mentioned curiosity and interest as two "crucial mentalistic concepts" when trying to develop a unified learning theory. Other such concepts are attention, motivation, and metacognition. An understanding of these concepts may also solve the old paradox which Iran-Nejad et al. formulated into the question how it is possible that children learn so much before school and sometimes so little at school.

The use of terms *curiosity, interest,* and *intrinsic motivation* vary, and sometimes they are used as synonyms with almost equivalent meanings. The aim of the present paper is not to invent some new concept or idea of curiosity, interest, or intrinsic motivation, nor even to specify what people ought to mean by these terms. The objective is rather to study, clarify, compare, and more thoroughly understand them. My intention is to examine these educationally interesting concepts and describe, if possible, how they overlap, what similarities they show to each other, and how they differ. I begin with concept of curiosity, and my starting point is that the meaning of a term is its use (cf. Wittgenstein, 1981).

Curiosity

The word curiosity is also used in everyday language, where it has, both in Finnish and in English (cf. Berlyne, 1978; Voss & Keller 1983, pp. 1-5), a slightly negative connotation. It is something that is not very desirable: it is not a necessary component of good education. Usually the word curiosity creates mental images of a person who is “a eavesdropper”, or "a nosy parker" , and thus has an excessive and indelicate desire to know something which is in some way

BEST COPY AVAILABLE

secret and non of his or her business. In this particular context Berlyne (1978, p. 99) used the word *inquisitive* as a synonym to *curious*.

When the word curiosity is used as a scientific concept, it does not have a negative connotation. However, it has been very problematic to define this usage. Fowler (1965, p. 23) expressed this problem by writing that curiosity is "a behaviour without a definition". The use of curiosity as a technical term has also been ambiguous. Curiosity has been used both as a hypothetical construct (e.g., Berlyne, 1960) as well as a concept which stresses the observable behavioral aspect (e.g., Penny, 1965). Exploration, on the other hand, has also been used both as a theoretical construct and as a description of an observable behavior which follows curiosity. In order to avoid this confusing situation, Voss and Keller (1983, p. 150) recommended that "*exploration* should be used to describe observable behavior, whereas *curiosity* should be used to describe the corresponding hypothetical construct". Moreover, it is noteworthy that in psychoanalytic literature curiosity has a slightly sexual connotation (see, e.g., Voss & Keller, 1983, pp. 3-4; Miller, 1987). Because the conceptualization of curiosity and many other related concepts, e.g. interest and intrinsic motivation, was greatly influenced by the work of Daniel Berlyne, his research on curiosity is discussed in some detail first.

State-like Specific Curiosity

Berlyne approached curiosity from the neobehavioristic point of view (see Berlyne, 1975; Madsen, 1974). He defined curiosity as an internal state occasioned when a person is confronted by object or event which is, for instance, novel or complex. Then the person undergoes an aversive state of subjective uncertainty which generates a tendency to engage in exploratory or investigatory behavior aimed at resolving or mitigating the uncertainty. This motivation is what Berlyne (1960, 1978) meant by *curiosity*, and the behavior is *specific* exploration. Berlyne's definition was based on the theory of optimal arousal first presented by Hebb (1955) and Leuba (1955), according to which a person has an optimal level of arousal or activation which he tries to maintain. This level of arousal is controlled by extrinsic and intrinsic variables. Berlyne used the term "collative properties" or "collative variables" when referring to stimulus characters which cause an interaction between the observer and the stimuli. In this process the perceiver collates or compares his or her mental schemata to the stimulus and the comparison results in a relative assessment of novelty, surprisingness, complexity, ambiguity, incongruity, and other properties which contain a certain measure of unexpectedness and uncertainty. Thus, collative variables involve conflict, and therefore competing and mutually interfering response tendencies that heighten arousal and lead to an aversive internal state which Berlyne called curiosity. The tendency to engage in exploratory behavior is a result of curiosity. The aim of this behavior is to reduce this uncertainty and in this way to recover the state of optimal arousal. Most of the modern definitions of curiosity depend on the ideas of Berlyne (e.g., Keller, Schneider, & Henderson, 1994). Because a state of subjective uncertainty is also a problem-solving situation, the concepts of competence, mastery, or effectance motivation have also been associated with concept curiosity (e.g., Trudewind & Schneider, 1994; White, 1959).

Berlyne (1960) made a distinction between *perceptual* and *epistemic* curiosity. Uncertainty-relieving perceptions activate perceptual curiosity and exploratory behavior. Berlyne (1963) stated that the degree of response conflict depend upon the nature of previous experiences with similar stimulation. Moreover, it depends on the number of opposing response tendencies, their degree of opposition, their relative strengths, and their absolute strength. According to Day and Berlyne (1971) the last factor is "somehow related to how meaningful the whole situation is to the observer, for the more meaningful, the greater is the intensity of conflict" (p. 313). If perceptual curiosity is aroused a person tries to resolve the conflict by specific exploration which can take the form of *receptor adjusting, locomotory exploration,* or *investigatory behavior.* The first happens, for instance, when a person comes into a room and fixes his or her attention on a specific object. Locomotor exploration happens when the person moves toward the source of stimulation. Investigator behavior is mainly manipulative, handling a strange object, taking parts away from it, or similar actions. Often an exploratory response includes more than one type of behavior.

Epistemic curiosity results from *conceptual conflict,* by which Berlyne (1963, 1978) meant conflict due to discrepant thoughts or beliefs or attitudes. The types of conceptual conflict are, for instance, doubt, perplexity, contradiction, conceptual incongruity, confusion, and irrelevance. By the term "epistemic behavior" Berlyne referred to behavior whose function is to get information that can relieve or mitigate the conceptual conflict. Berlyne noticed that in exploratory behavior a person deals mainly with the perception of objects or events, whereas in epistemic behavior one deals with concepts and symbolic representations. Moreover, the function of exploratory behavior is to provide stimuli that will be immediately useful, whereas the function of epistemic behavior is to "equip the organism with knowledge," by which Berlyne meant "structures of symbolic responses."

Berlyne (1965) divided epistemic behavior into three categories, namely, epistemic *observation*, which includes different kind of experimental and observational techniques, *consultation*, which includes asking other people questions or consulting books, and *directed thinking*. In everyday life, exploratory and epistemic behavior can frequently be intertwined, for instance, when a child leafs through an animal book and sees for the first time a picture of a kiwi. The child stops scanning and fixes his or her attention on the picture, after that the child ask questions like "What is this?" "Is this a bird or some other creature?" first to him- or herself and then maybe to a parent. If the parent cannot answer the child may pose the same questions to a teacher the following day.

The Broad Meaning of Curiosity

Day (1968) later extended Berlyne's definition of curiosity. To clarify curiosity construct, Berlyne (1960) had divided exploration into *specific* and *diversive* exploration. Berlyne had stated that what he means by curiosity is always specific, which means that the exploration that follows curiosity is always specific exploration. The aim of this specific exploration is to release the curious person from the subjective uncertainty caused by collative variables. Analogically Day suggested that curiosity can also be dichotomized to *specific* and *diversive curiosity*. According to Day, diversive curiosity is the condition of heightened arousal induced by a situation of changelessness, repetition, or monotony. Day and Berlyne (1971) specified that diversive curiosity also results from uncertainty, but leads to diversive exploration such as seeking entertainment or new experiences. Moreover, Hutt (1981) noticed that diversive exploration also includes playful behavior. Thus, the aim of exploration is not to reduce uncertainty but to increase the level of activation or to provide stimulation. That is, as Day and Berlyne (1971, p. 312) expressed it, "looking for collative variability." Day (1968) noted that what he means by diversive curiosity may be analogous to what Maw and Maw (1965) considered as "the need to seek new experiences" or to extend one's knowledge into the unknown.

In the definition of diversive curiosity presented by Day (1968, 1971), the level of optimal arousal is approached from its opposite, namely from specific curiosity. If the state of arousal, resulting perhaps from boredom, has fallen below the optimal level of arousal, then the result may be diversive curiosity. As Boyle (1983, p. 380) indicated, the resultant behavior depends on a person's cognitive appraisal of a low stimulus situation which may induce a psychological state of either diversive curiosity or anxiety. The diversively curious person tries consciously look for new, amusing, or exciting stimuli in order to raise the level of arousal to the optimal plane. Exploration cannot be specific and diversive at the same time, but specific and diversive curiosity may temporarily form a sequence (cf. Hutt, 1970).

Zuckerman (1971, 1979, 1984, 1987) used the term "sensation seeking" and Spielberger and Starr (1994) term "experience seeking" in very much the same way as Day used the term "diversive curiosity" (i.e., to refer to a desire or need to experience sensory stimulation). Zuckerman used term "sensation seeking" to describe a personality trait which shows itself in various forms of sensation seeking. Zuckerman used the term "sensation" instead of "stimulation" because he emphasizes the role of the subjective element of the stimulus; the same stimulus may product different sensory experiences and emotions in different people. According to Zuckerman (1979, p. 10) "it is the sensory effects of external stimulation that are most important." In a factor analytical study (Zuckerman, 1971) sensation seeking was sharpened to four factors: "thrill and adventure seeking," "experience seeking," "disinhibition," and "boredom susceptibility." Zuckerman's theory of sensation seeking utilizes the modern idea of optimal level(s) of arousal construct (see Zuckerman, 1984, 1987).

Day (1971, p. 102) also extended the concept presented by Berlyne in another respect. Berlyne described curiosity as a state. In addition to this, Day invoked a "trait characteristic of curiosity" (i.e., curiosity motive as a behavioral disposition). According to the development of trait theories this difference has been ascribed to a view of curiosity both as a motive-like trait (C-trait) and state (C-state) (see Naylor, 1981; Boyle, 1983, 1989). For example, Boyle (1983) proposed a schematic conceptual model of state-trait curiosity in which the cognitive appraisal of external or internal stimuli plays a central role. According to the state-trait distinction individual differences in exploratory actives have been conceived to vary along two dimensions: (a) the trait of the individual, referring to the predisposition to manifest a state across a wide range of contexts and conditions; and (b) the state of the individual, referring to affective reactions that vary in intensity, fluctuate over time, and result from specific environmental conditions and level of the trait that an individual possesses. It is also presumed that those possessing more C-trait experience greater intensity of C-state. Trudewind and Schneider (1994) described the relationship between C-trait and C-state as follows:

> we postulate an original motive to explore one's physical and social environment. The ultimate function of this behavioral disposition is the acquisition of

> knowledge or the assimilation of objective structures, whereas the immediate cause of exploration is assumed to be a state of subjective uncertainty created by certain aspects of the environment. (p. 152)

Wohlwill (1981) reformulated the Berlyne-Day "specific-diversive" distinction with the terms *inspective* and *affective*, since both concern responses to specific stimuli such as inspection for the sake of uncertainty or conflict-reduction, or contemplation for the sake of enjoyment or pleasure. The inspective-affective distinction is no longer a mutually exclusive dichotomy, as was the specific-diversive distinction (see Naylor, 1981). Instead of perceiving it as a dichotomy, Wohlwill perceived the inspective-affective differentiation as a "continuum" (cf. Nunnally, 1981). Inspective and affective exploration are closely interdependent, and in practice exploration usually contains varying mixtures of these two. Wohlwill (1987, p. 64) emphasized that "exploration of a stimulus may (and generally does) serve both an information-extraction and affect-production function." Moreover, Keller, et al. (1994) noticed that both anticipated positive emotions and previously experienced ones have an influence on exploration. Wohlwill (1987) later expanded the inspective-affective differentiation by adding the *genuine diversive exploration* concept to his conceptual system. Wohlwill described diversive exploration as an activity in search of stimulation, designed to relieve boredom, raise arousal, or the like. Voss (1987, p. 47) summarized Wohlwill's position as follows: "Whereas the functions of inspective and diversive exploration are uncertainty reduction and stimulus/sensation seeking, respectively, affective exploration is directed to the maintenance of an optimal hedonic tone."

Among others Berlyne (1960), Boyle (1983), Keller (1987), Schneider and Unzner (1994), and Spielberger and Starr (1994) noticed that curiosity-instigating collative variables also arouse fear, neophobia, and anxiety at the same time. Curiosity and anxiety are antagonistic, and exploration is a compromise between both behavioral tendencies - curiosity, which is reduced only by inspecting, and anxiety, which leads to withdrawals or avoidance behavior. Spielberger and Starr (1994) called theories which explicate incompatible exploratory and avoidance reaction to collative variables *dual-process* theories of curiosity and anxiety. These theories point out that when diversive curiosity (or sensation-seeking) is strong and anxiety is relative weak, diversive exploration is motivated. On the other hand, when the anxiety is much stronger than diversive curiosity, avoidance behavior (flight reactions) will occur. However, the simultaneous experience of curiosity and anxiety seems to be symbiotic in motivating specific exploratory behavior when the physiological arousal associated with these emotional states is at or near an optimal level. According to Spielberger and Starr (1994, p. 233), at this level the "reduction in collative stimulus intensity will increase the individual's overall experience of pleasantness by reducing the unpleasantness associated with moderately high levels of anxiety, while the pleasantness associated with high curiosity remains unchanged." Thus, diversive and specific curiosity can be explained in terms of the intensity of curiosity and anxiety as emotional states. According to this view, which Spielberger and Starr called the "Optimal Stimulation/Dual Process Theory of Exploratory Behaviour," the concept of specific curiosity drive appears to be redundant although it still is important to distinguish between diversive and specific exploration.

The scope of the concept of curiosity varies. Many factor analytical studies support the claim that curiosity is not a unitary construct (e.g., Ainley, 1987; Byman, 1993; Kreitler, Zigler & Kreitler, 1975; Langevin, 1971; Olson & Camp, 1984; Pearson, 1970; Rubenstein, 1986). But we do not yet know how many "components" curiosity has. In this article I have focused my attention only on the main meanings of the term "curiosity." Thus, I have mentioned only those distinctions which seem to be the most important, especially from the viewpoint of school motivation.

Interest

In everyday language interest has many more positive connotations than curiosity: it is even desirable. Commonly interest is used as an descriptive term but also as an everyday explanation for learning (e.g., Paul learns X very easily because he is interested in X). However, the scientific concept interest is hardly the same as its common-language everyday counterpart. According to Valsiner (1992, p. 30) "'interest' in our analytic schemes is unlikely to be the same as the sense in which 'interest' as an everyday concept is used by persons in their life-worlds. Because the term is the same, however, it is easy to succumb to the illusion that our meaning for it equals the personal sense of it in the minds of our subjects." Schiefele (1992) noticed that the most obvious problem confronting studies concerning interest is that the everyday meaning of the term interest is assumed to adequately delineate the concept. This view has led to very narrow measurements of interest, where only a single rating scale has been used to measure the interest which a topic or a sentence awakens.

The problem of conceptualizing interest is the same as with curiosity: it has been done in a variety of ways, based on different theoretical orientations and different research methods. However, as Iran-Nejad (1987) has noted interest has been treated as an affective variable, a general arousal experience, or an emotion. Piaget's cognitive-affective distinction has been the starting point of many studies of interest. Piaget (1981, p. 5) noticed that "affective states that have no cognitive elements are never seen, nor are behaviors found that are wholly cognitive." Thus, assimilation and accommodation also have both a cognitive and an affective side. The affective aspect of assimilation is interest, and the cognitive aspect is understanding. According to Piaget "accommodation in its affective aspect is interest in the object in as much as it is new. In its cognitive aspect, accommodation is the adjustment of schemes of thought to phenomena" (p. 5).

Izard (1991) pointed out that it is only the emotional component of the affective-cognitive structure or feeling-thought bond that provides motivation. Izard determined ten fundamental emotions, one of which is interest. Moreover, she noticed that interest is the most frequently experienced positive emotion which also motivates the development of skills, competencies, and intelligence. At the experimental level Izard described interest-excitement as the feeling of being engaged, caught up, fascinated, or curious. According to Izard "there is a feeling of wanting to investigate, become involved, or expand the self by incorporating new information and having new experiences with the person or object that has stimulated the interest. In intense interest or excitement the person feels animated and enlivened" (p. 100). Izard's description of the subjective part of intense interest comes very near to that state of which Csikszentmihalyi (1975, 1990) labelled "flow." Csikszentmihalyi determines flow as a state in which people are so involved in an activity that everything else goes out of mind, the experience itself is enjoyable and so the only reward seems to be the ongoing subjective experience of enjoying the activity.

According to Krapp, Hidi, and Renninger (1992, p. 5) interest is most commonly defined as "a phenomenon that emerges from an individual's interaction with his or her environment". It may be either short or enduring (persistent). Krapp et al. have distinguished two major points of view from which interest has been approached. One is interest as a characteristic of person and the other is interest as a psychological state aroused by specific characteristics of the learning environment. Traditionally, the former approach has been termed with term "individual interest" or "topic interest" and the latter has been called "situational interest." According to Hidi (1990) individual interest develops slowly and tends to have long-lasting effects on a person's knowledge and values, whereas situational interest is an emotional state that is evoked suddenly by something in the immediate environment and that may have only a short term effect on an individual's knowledge and values. However, individual and situational interest are not dichotomous phenomena. Both types of interest concern person/environment interaction, and they are supposed to interact and influence each other's development.

Schiefele (1991) drew a conceptual distinction between a *latent* (disposition) and an *actualized* individual interest. A latent individual interest is a relatively enduring preference for certain topic, subject areas, tasks, contexts, or activities. Moreover, Schiefele suggested that interest is a content-specific concept as well as a directive force, and that it consists of two kind of valences: feeling-related and value-related valences. Feeling-related of valences are feelings that are associated with a topic or an object, for instance feelings of enjoyment and involvement. Value-related valences refer to the attribution of personal significance to an object. Thus, some objects of interest are preferred because involvement with them creates, for instance, strong feelings of excitement, whereas other objects of interest are preferred because they may have high personal relevance. According to his valence distinction Schiefele (1992, p. 154) reinterpreted interest "as a domain-specific or topic-specific motivational characteristic of personality, which is composed of intrinsic feeling-related and value-related valences." Schiefele also emphasized that interest has an intrinsic character which he termed "self-intentionality." According to Schiefele, for instance, "in the context of text learning, this means that the learner should be involved in a topic for its own sake and not for any external reason (e.g., passing an exam)" (p. 155). Schiefele (1991) described actualized individual interest as "a content-specific intrinsic motivational orientation," which means that a person is in a state of being interested in a certain topic and wants to learn about (or become involved with) that topic for its own sake.

Unlike individual interests, which are always specific to individuals, situational interest is assumed to be spontaneous, fleeting, and shared among individuals. Situational interest may be evoked suddenly by such aspects of a situation as novelty, intensity, or complexity. If these aspects are seen as characteristics of an environment, then they contribute to the *interestingness* of the situation. Moreover, Deci (1992) emphasized that a situation or an object must also offer an "optimal challenge" in order to be interesting. However, situational interest can also be seen as a psychological state within the person. Hidi (1990) investigated a specific form of situational interest which she labelled "text-based interest." It results from the interaction of textual features and the individual who is reading the text. Hidi found out that two types of factors contribute to text-based interest. One group of factors was what Berlyne called collative variables (i.e., variables that create uncertainty in the subject). The second group of factors consists of universally interesting concepts (e.g., murder and sex), human activity, intensity factors, and life themes.

Hidi and Anderson (1992) tried explicate the conceptual distinction between two psychological states of interest: actualized individual interest and situational interest. They hypothesize that the difference is in the area of affect. Liking and interest have been connected in many investigations. Asher (1979, p. 687), for instance, crystallized the definition of individual interest to the expression "something is interesting when you like it and would like to find out more about it." Moreover, Deci (1992) saw that interest, excitement, and enjoyment are only "slightly different experiences." Thus, feelings such as liking, enjoyment, and involvement are typically connected to actualized individual interest. However, the connection of situational interest and liking (or pleasingness) seems not to be so simple. Berlyne (1971a, p. 217) first suggested that "interestingness may continue to rise, while pleasingness sharply declines, when moderate degrees of complexity are exceeded." Iran-Nejad (1987, p. 121) later referred to the same thing by noting that "a snake can be interesting without being liked, and a particular soft drink may be liked without being interesting." Based on these arguments Iran-Nejad (1987) argued that the underlying causes of interest and affect are different. Thus, intense intellectual activity can also cause intense interest.

"Interest" and "Curiosity"

The things that arouse curiosity also arouse interest. Persistence is a fundamental characteristic of both curiosity and interest. Exploration follows both curiosity and interest. Moreover, anxiety has been reported to have a negative connection to both curiosity and interest. Thus, it seems difficult to differentiate between the concepts curiosity and interest. For example, Day (1982, p. 19) defined curiosity with the help of the term interest. According to Day "curiosity is a state of excitement and directed interest." Moreover, Langevin (1971) noticed that two curiosity factors that he labelled "breadth of interest" and "depth of interest" may be the same which Berlyne and Day conceptualized as diversive and specific curiosity. Following Langevin, Ainley (1987) used both terms "curiosity" and "interest" when defining two styles of behavior, namely breadth-of-interest and depth-of-interest curiosity. Maw (1971) suggested that interest is "less mobile and more positive than curiosity." Haywood (1971, p. 113) sharpened this distinction, proposing that curiosity, while similar to an inferred state of the organism, can be inferred more readily from observation of behavior than can interest. According to Haywood curiosity would seem to be the more scientifically useful construct of the two.

According to Hidi (1990) the same variables that stimulate curiosity can also stimulate situational interest (i.e., so called "collative variables"). The underlying characteristic of these variables is that they create conflict and uncertainty (see Berlyne, 1978). However, Hidi and Anderson (1992) later tried to make explicit the differences between situational interest and curiosity. According to Hidi and Anderson the most important differences are: (a) situational interest can be elicited not only by collative variables, but also by content-specific text characteristics such as power, death, and sex; (b) situational interest may develop into relative enduring individual interests; and (c) the inverted-U function does not necessarily characterize the relation between situational interest and the stimulus characteristics that elicit it.

Tobias (1994) concluded that the curiosity construct has two advantages over interest. The first is that curiosity can be related to three apparently different states which according to Tobias are "an eagerness to approach some activities and situations motivated by curiosity and interest, neutral reactions, and disinterest leading to flight induced by anxiety from other material" (p. 47). The second advantage of the curiosity construct is that a number of measures are available for research. At this connection Tobias referred to curiosity in its broad meaning including sensation-seeking, novelty experiencing, academic curiosity, curiosity as an state, and curiosity as a trait (cf. Ainley, 1987; Byman, 1993; Langevin, 1971).

Harty, Samuel and Beall (1986) demonstrated empirically the close relationship of the following concepts: attitudes toward science, interest in science, and science curiosity. In an factor analytical study the all three attributes had high loadings on the same factor. Thus, Harty, et al. came to the conclusion that the attributes of attitudes toward science, interest in science, and science curiosity might be similar and might be a single construct; or at least that the items of the three employed instruments are measuring the same factor.

Kirkland (1976a) presented a sequential model that interlocks four concepts: attention, curiosity, skill, and interest. The sequence begins when curiosity is triggered by "attention to an environmental anomaly." Sustained and persistent effort applied to resolve curiosity leads to skill development. Kirkland defines interest as the voluntary application of a skill. It develops from successful resolution of puzzles. Day and Maynes (1972) specified that curiosity can contribute to the development of interest, especially for example scientific and mechanical interest. Kirkland (1976b) also suggested that established interest can direct previously aroused epistemic curiosity (i.e., the curiosity reduced by rehearsal of knowledge).

Recently Krapp (1994) advanced a theory which links the two concepts diversive curiosity and individual interest. Krapp emphasized that "diversive curiosity and exploration are not directed randomly at whatever objects or action possibilities happen to be available, but instead often exhibit a goal-oriented character. According to this view, interests are an important component of this phase of stimulus-search behavior"(p. 96). In addition, interests also play a "decisive role" in the content orientation of specific curiosity. Loewenstein (1994, p. 93) emphasized that interest "primes the pump" of specific epistemic state curiosity which means that a person's pre-existing interests focus attention and in this way also effects their curiosity. Loewenstein interpreted epistemic curiosity as a "feeling of deprivation" that results from an awareness of an information gap. By information gap Loewenstein meant a "discrepancy between what one knows and what one wishes to know". This view implies that curiosity requires a pre-existing knowledge base because without information there can not be an information-gap. Thus, at first it may be necessary to stimulate information gathering without curiosity, for instance using extrinsic rewards. Curiosity can occur when the person's "informational reference point" has been elevated to a certain level.

Intrinsic motivation

Unlike the terms "curiosity" and "interest," the term "intrinsic motivation" is not used in everyday language. However, Harter (1981, p. 309) has observed that the term "intrinsic motivation" has been used "rather loosely" in scientific discussion. The concept is derived from Woodworth's (1918, 1958) "behavior primacy theory," according to which motivation consists fundamentally of dealing actively with the environment. As Madsen (1974, p. 424) noticed, this theory did not dismiss "'extrinsic' motivation". However, the main idea was that even without extrinsic sources of motivation the organism would be active. According to Deci's (1992) interpretation the concept intrinsic motivation emerged from the critique of Skinnerian operant theory and Hullian drive theory. These theories where not adequate to explain such activities as exploring novel spaces and manipulating objects, actions that seemed to have neither a direct nor an indirect relation to reinforcements. As a result of the attacks against Skinner's functional theory of reinforcement, intrinsically motivated behaviors were defined as those that occur in the absence of any operationally separable reinforcement. Deci pointed out that this view has also served as an *operational definition* of the concept for the past two decades. On the other hand, intrinsically motivated behaviors were defined to refute Hull's drive theory statement that all behavior is just as reducible to *physiological drives* as those specific behaviors that are based on the innate *psychological needs* of the organism. At this connection Deci (1992) emphasized the significance of two inherent psychological needs: self-determination and competence. By self-determination Deci meant that a person involves in an activity "with a full sense of wanting, choosing, and personal endorsement." Intrinsically motivated behaviors are always self-determined.

Originally Deci (1975) split motivation into intrinsic and extrinsic motivation. Extrinsically motivated behavior is instrumental in nature. Such action are performed for the sake of some expected outcome or extrinsic reward or in order to comply with a demand. Intrinsically motivated behaviors, on the other hand, are engaged in, as Deci expressed it, "for their own sake and not because they lead to an extrinsic reward." Later Deci and Ryan (1985) limited the idea of the antagonistic nature of intrinsic and extrinsic motivation. Borrowing the concept of internalization from Schafer (1968), Deci and Ryan (1985) identified four types of extrinsic motivation: external, introjected, identified, and integrated forms of regulation. More recently, Rigby, Deci, Patrick and Ryan (1992) stated that the simple intrinsic-extrinsic dichotomy is also insufficient for a full understanding of the motivational processes that underlie learning. Moreover, Harter and Jackson (1992) demonstrated that intrinsic-extrinsic motivation must be conceptualized both as a trait and a nontrait in order to fully understand children's motivational orientation to school subjects.

According to Deci and Ryan (1985, pp. 32-35) intrinsic motivation is based in the innate or psychological, organismic needs for competence and self-determination. Deci, Vallerand, Pelletier, and Ryan (1991) added to the list a need for relatedness, in other words a need to develop secure and satisfying connections with others in one's social milieu. They also used the term "autonomy" as a synonym of the term "self-determination." Deci (1975) used the concept of competence in the same way as White (1959) in his theory of "effectance motivation." By competence White referred to "an organism's capacity to interact effectively with its environment." Deci et al. (1991) compressed the idea into single thought by saying that "competence involves understanding how to attain various external and internal outcomes and being efficacious in performing the requisite actions." According to Deci and Ryan (1985) needs for competence and self-determination motivate people to an ongoing process of seeking and attempting to conquer "optimal challenges." Moreover, according to Deci (1992, p.51) "where there is interest, there is likely to be optimal challenge or novelty."

Deci and Ryan (1985) utilized the concept autotelic when they explained the teleology of intrinsic motivation. Csikszentmihalyi (1975) first used this concept when emphasizing the role of enjoyment or the inherent experiential aspects of intrinsically motivated behaviour: the reward is the ongoing subjective experience of enjoying the activity. Deci and Ryan (1985, p. 34) used this thought when they declared that "the emotions of enjoyment and excitement accompanying the experiences of competence and autonomy represent the rewards for intrinsically motivated behavior."

Harter (1981) leaned heavily on the distinction to intrinsic and extrinsic motivation when she spoke of intrinsic versus extrinsic orientation of school children. Harter described these orientations with five bipolar dimensions. In one of these five dimensions "learning motivated by curiosity" is the intrinsic pole and the extrinsic pole is "learning in order to please the teacher." The other dimensions in Harter's analysis are "incentive to work for one's own satisfaction" versus "working to please the teacher and get good grades", "preference for challenging work" versus "preference for easy work", "desire to work independently" versus "dependence on the teacher for help", and "internal criteria for success or failure" versus external criteria (e.g., grades, teacher feedback) "to determine success or failure." Like Deci Harter also used the concept of competence, originally represented by White (1959), to support her view of intrinsic motivation.

"Curiosity" and "Intrinsic Motivation"

The relationship between curiosity and intrinsic motivation will depend on which conceptualization one chooses. According to Beswick (1974) curiosity is "the prototypical example" of intrinsic motivation (see also Bruner, 1966). Intrinsic motivation has also been used as a synonym for curiosity (e.g., Rubenstein, 1986). Hunt (1971) defined intrinsic motivation as follows:

> By *intrinsic motivation* I mean that motivation which is inherent in information processing and action the notion of intrinsic motivation emphasizes the role of cognition in both affection and conation It emphasizes, in other words, what Professor Berlyne has termed 'collative variables' (p 1)

Day and Berlyne (1971, p. 334) determined that curiosity or intrinsic motivation is "a state of tension arising from response conflict due to uncertainty or insufficient knowledge about a percept or concept that leads to further study or exploration, with the goal of obtaining additional information and relieving the conflict and tension."

Usually the concept intrinsic motivation has a broader meaning than curiosity, and curiosity is only one component of intrinsic motivation (e.g., Berlyne, 1971b; Harter, 1981; Gottfried, 1985). Deci (1975, p. 54), referring to the work of Berlyne, summarized by saying that "all uncertainty is not aversive; in fact, organisms sometimes seek uncertainty. Further, the need to reduce uncertainty does not provide a general theory of intrinsic motivation." Thus, what Deci meant by intrinsic motivation is somehow broader than only specific curiosity. On the other hand, if the concept curiosity is used in its broad meaning containing both specific and diversive curiosity, then the concepts intrinsic motivation and curiosity seem to be almost identical. This conclusion is also implicit in Deci's theory. According to Deci (1975) there are two general kinds of intrinsically motivated behaviors. First, "when there is no stimulation people will *seek* it" (p. 61) or more detailed, people seek out challenges which are optimal for them. The other type of intrinsically motivated behavior involves, according to Deci, "*conquering* challenge or reducing incongruity." On the other hand, only when a person is successful in reducing incongruity which he or she meets or creates will that person feel competent and self-determining. Thus, more than Day and Berlyne, Deci emphasized the role of competence and self-determination in intrinsic motivation.

More recently Vallerand, Pelletier, Blais, Brière, Senècal, and Vallières (1992) have split intrinsic motivation (IM) into three types: IM to know, to accomplish things, and to experience stimulation. Intrinsic motivation to know can be defined as the performing of an activity for the pleasure and satisfaction that one experiences while learning, exploring, or trying to understand something new. IM to know relates to such concepts as exploration, curiosity, learning goals, intrinsic intellectuality, IM to learn, and to more global concepts such as that of the epistemic need to know and understand, and that of the search for meaning. The concept need for cognition presented by Cacioppo and Petty (1982) may also be added to the list (cf. Olson, Camp & Fuller, 1984; Tolentino, Curry & Leak, 1990).

The second type of IM which Vallerand et al. (1992) have called intrinsic motivation to accomplishments comes very near to the concept of mastery motivation used by Harter (1981). In his theory of effectance motivation White

(1959) was the first to state that individuals interact with the environment in order to feel competent and to create unique accomplishment. According to this view IM to accomplish things can be defined as engaging in an activity for the pleasure and satisfaction experienced when one attempts to accomplish or create something. The third type of intrinsic motivation, IM to experience stimulation, relates to activity to experience stimulating sensations derived from one's engagement in the activity. As an example of this kind of stimulating sensations Vallerand et al. (1992) mentioned sensory pleasure, aesthetic experiences, as well as fun and excitement. By the concept intrinsic motivation to experience stimulation, Vallerand et al. (1992) seem to mean much the same as Zuckerman's (1971, 1979, 1984) sensation seeking and Day's (1968, 1971) concept of diversive curiosity. Thus, the definition of intrinsic motivation presented by Vallerand et al. (1992) comes very close to the broad definition of the concept curiosity (cf. Ainley, 1987; Byman, 1993; Langevin, 1971, 1976).

Gottfried (1985, 1990) used the term "academic intrinsic motivation" in a broad sense to depict a special kind of intrinsic motivation for school learning. Academic intrinsic motivation involves enjoyment of school learning characterized by a mastery orientation; curiosity, persistence, taskendogeny, and the learning of challenging, difficult, and novel tasks. Gottfried never exactly explains what she meant by the term "curiosity" in her intrinsic motivation construct but she refers to Berlyne, which may mean that the curiosity she means is specific and not diversive.

"Interest" and "Intrinsic motivation"

As has already been noted, intrinsic motivation is a purely scientific term and has no common-language meaning. However, many of those phenomena that psychology explain by intrinsic motivation are commonly referred to under the label of "interest". Schiefele (1991) described the relationship between the terms "interest" and "intrinsic motivation" as follows:

> To many psychologists, interest is a vague, everyday term that denotes a personal characteristic or an affective state and that has already been thoroughly investigated by modern motivational psychology. Specially, it seems as if interest is nothing more than the lay term for intrinsic motivation. There is some reason to believe, however, that intrinsic motivation research does not capture all of the essential aspects of interest. (p. 299)

Thus, the terms intrinsic motivation and interest have been used interchangeably.

Intrinsically motivated acts are performed freely, voluntarily and without the necessity of material rewards or constraints. According to Deci (1992) intrinsically motivated behavior is characterized by "concentration and engagement, it occurs spontaneously and people become wholly absorbed in it." In this connection Deci referred to Csikszentmihalyi's concept of flow. Deci argued that the state of flow also can be though of as a prototype of being interested. Moreover, Deci (1992, pp. 48-49) admitted that it is difficult to talk about intrinsically motivated activity without describing people being interested in the activity. The so-called emotions "interest", "excitement," and "enjoyment" are distinctive marks of intrinsically motivated activity. Thus, Deci and Ryan (1985, p. 29) stated that interest and excitement are central emotions that accompany intrinsic motivation, and the concept of flow describes some of the purer experiential aspects of intrinsic motivation. When people are highly intrinsically motivated, they will also be extremely interested in what they are doing and experience a sense of flow. Moreover, Snow and Jackson (1994) have concluded that interest is an emotional outcome or reward of intrinsic motivation. Thus, it is obvious that it is impossible to determine the subjective part of intrinsic motivation without the help of the concept of interest. Deci (1992, p. 55) also related interest to self-determined extrinsic motivation. Deci suggested that "a person can gradually become self-determined with respect to the activity even though he or she may not be very interested in it. It is probable that, as the regulation becomes integrated, the person will experience greater interest in doing it." However, Deci noticed that importance may be more important than interest in describing to self-determined extrinsic motivation.

Concluding Remarks

In this article my aim has been to clarify and more thoroughly understand three educationally interesting and related concepts, namely curiosity, interest, and intrinsic motivation. All three of these concepts have been claimed to have a positive effect, for instance, on learning. My starting point was that the meaning of a term is its use. In this sense I analyzed the common-language meanings of words "curiosity" and "interest." The major difference in the use of these words is that the connotation of the term "interest" is much more positive than that of the term "curiosity". Reference to "interest" is also common for teachers when they try to explain good or bad learning results of their students. Thus, interest has a important role also in teachers' implicit or subjective learning theories.

BEST COPY AVAILABLE

Curiosity, on other hand, is not necessarily a desired trait in every any pupil because it invokes many negative images in the minds of people; a curious person is "too interested" and usually even in wrong things. For a child "natural curiosity" is allowed, but not for an adult. The fact that both curiosity and interest have both a common-language and a scientific meaning is problematic. Valsiner (1992) refers in this connection to the "confusion of levels." This confusion of levels may be especially disastrous for the concept of curiosity because the common-language and scientific connotations of this word are completely different.

One of my aims was to discover to what extent the concepts intrinsic motivation, curiosity, and interest overlap, what family resemblance they show, how they differ, and how they are related. Answering these questions is difficult because often scholars do not define their terms carefully. Usually only the main concept is described thoroughly, and the others are used as if they were self-evident. Paradigm differences are the other problem of conceptual elaboration: different paradigms have different premises. Behaviorism and neobehaviorism have strongly effected, for instance, the conceptualization of curiosity. Because of this, very little is known about the affective qualities of different levels of subjective uncertainty and curiosity (cf. Keller et al. 1994). The only noteworthy work in this area is Wohlwill's (1981, 1987) reformulation of exploration away from a specific-diversive distinction (see, e.g., Day & Berlyne, 1971) in the direction of an inspective-affective distinction. What Wohlwill however really reformulated with regard to specific exploration seems to be only the other part of this exploration, namely the exploration that is induced by perceptual curiosity. How well the inspective-affective distinction is suited to an understanding of the specific exploration aroused by epistemic curiosity needs further clarification.

The scientific concepts of curiosity, interest, and intrinsic motivation overlap. For instance, some researchers use the term "curiosity" and the term "intrinsic motivation" interchangeably, and this is a good reason for concluding that the meanings of the two terms overlap. However, the range of meaning of the term curiosity differs a great deal. In general, the term "curiosity" may be substituted for "intrinsic motivation" in any context where we use the term curiosity in the sense Day (1971) used the term, thus meaning both specific and diversive curiosity; however that substitution cannot be made with equal ease when we mean by curiosity what Berlyne (1960, 1978) termed specific curiosity. In this meaning the term curiosity refers to a subjective state of uncertainty which a person usually wants to mitigate or resolve by seeking information, whereas the concept of intrinsic motivation is somewhat broader and includes, for instance, both the creation and resolution subjective uncertainty. Moreover, the needs for competence and self-determination are supposed to emphatically underlie intrinsic motivation.

Of the three concepts intrinsic motivation, curiosity, and interest, it seems the most difficult to differentiate curiosity from interest, or more detailed state-like specific curiosity from situational interest. Hidi and Anderson (1992) made a good attempt to differentiate between these concepts although it is little unclear whether they mean by curiosity only specific curiosity (see Berlyne, 1960, 1978). Nevertheless, the three distinctions made by Hidi and Anderson need further clarification. The first proposition that collative variables elicit both situational interest and curiosity, but such concepts as death and sex or more commonly, emotionally loaded life themes to which a person can identify elicit only situational interest but not curiosity, is tightly bound to the conceptualization of curiosity we prefer. Freud, for instance, hypothesized that later curiosity originates from infantile sexual impulses (see Miller, 1987). Moreover, if by curiosity we also mean diversive curiosity (sensation seeking) the difference between state curiosity and situational interest as proposed by Hidi and Anderson becomes even more fuzzy. The second proposition that the inverted-U function does not necessarily characterize the relationship between situational interest and the stimulus characteristics that elicit it is questionable in the sense that this relationship does not necessarily describe the relationship between curiosity and the stimulus characteristics that elicit it (see, e.g., Boykin & Harackiewicz, 1981). The third proposition of Hidi and Anderson seems useful, but needs specifying. It seems true that situational interest may develop into relative enduring individual interests. On the other hand, there are several studies (e.g., Boyle, 1983, 1989) which indicate that curiosity as a state and, curiosity a trait form are interactive psychological system and thus effect each others development. However, if Hidi and Anderson actually meant that individual interest and C-trait are different, they seem to be right. First, individual interest is not a personality trait (Schiefele, 1991). Second, C-trait is only a proneness or predisposition to manifest a C-state across a wide range of contexts and conditions, whereas individual interest is bound to a specific topic, object, or activity: interest is domain- or content-specific. Third, C-trait as a predisposition is a passive part of personality whereas individual interest is self-determined. This in turn, means that a person engages in activity with a full sense of wanting, choosing, and personal endorsement.

Thus, although Krapp et al. (1992, p. 9) argued that "conceptually, situational interest is similar but not identical to the concepts of curiosity and exploration" it seems difficult to distinguish these concepts by objective criterion.

However, as Scriven (1988) emphasized it is also important to analyze the subjective components of concepts. The subjective part of situational interest has been described by the word "pleasant," whereas state curiosity has been described typically with the word "aversive". However, this distinction needs additional clarification because it is a well known fact that people sometimes voluntarily seek curiosity invoking stimulus (e.g., mystery stories and crosswords). Thus, why do they seek curiosity if it is aversive? Boyle (1983) and later Keller et al. (1994) have suggested that, depending on the appraisal of the situation and/or individual differences, the balance of the aversive and positive motivational components following curiosity may vary. Moreover, it must be observed that, although curiosity itself is aversive, the process of satisfying curiosity is pleasurable. During that process a person can feel such positive feelings as competence and mastery (see, e.g., White, 1959; Day, 1982; Loewenstein, 1994).

One possible way to differentiate between curiosity from interest and intrinsic motivation seems to be to examine them on the externally-internally controlled dimension. This view also reflects the contradiction which is typical of behaviourism and cognitive psychology. The central premise of modern cognitive psychology is that human beings have active control over their psychological actions. According to this view the locus of control is inside the organism, whereas the central premise of behaviorism is that human activity is controlled by external stimulation. Interest and intrinsic motivation are mostly seen as self-determined and self-regulated (see, e.g., Deci, 1992; Krapp, 1994), whereas curiosity is typically seen as intensive and compelling thus possessing the basic characteristic of a primary drive (see Loewenstein, 1994; von Wright, 1986). However, as has been already noted depending on the conceptualization of curiosity there also exists voluntary curiosity.

Taken overall, the meanings of the terms "intrinsic motivation," "curiosity," and "interest" are closely related and difficult to differ rigorously. However, it is clear that at least the concept intrinsic motivation can not be fully understood without the help of the concepts curiosity and interest. Deci and Ryan (1985, p. 245), for example, noticed that "intrinsic motivation is in evidence whenever students' natural curiosity and interest energize their learning." According to the words of Wittgenstein (1980, 1981) intrinsic motivation is a family.

References

Ainley, M.D. (1987). The factor structure of curiosity measures: Breadth and depth of interest curiosity styles. *Australian Journal of Psychology, 39*(1), 53-59.

Asher, S.R. (1979). Influence of topic interest on black children's and white children's reading comprehension. *Child Development, 50*, 686-690.

Berlyne, D.E. (1960). *Conflict, arousal, and curiosity*. New York: McGraw - Hill.

Berlyne, D.E. (1963). Motivational problems raised by exploratory and epistemic behavior. In S. Koch (Ed.), *Psychology: A study of science (Vol.5)* (pp. 284-364).New York: McGraw-Hill.

Berlyne, D.E. (1965). Curiosity and education. In J.D. Krumboltz (Ed.). *Learning and the educational process* (pp. 67-89). Chicago: Rand McNally.

Berlyne, D.E. (1971a). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.

Berlyne, D.E. (1971b). What next? Concluding summary. In H.I. Day, D.E. Berlyne, & D.E. Hunt (Eds.), *Intrinsic motivation: A new direction in education* (pp. 186-196). Toronto: Holt, Rinehart & Winston.

Berlyne, D.E. (1975). Behaviorism? Cognitive theory? Humanistic psychology? To Hull with them all. *Canadian Psychology Review, 16*, 69-80.

Berlyne, D.E. (1978) Curiosity and learning. *Motivation and Emotion, 2*, 97-175.

Beswick, D.G. (1974). Intrinsic motivation in senior secondary students. *Education Research Perspectives, 1*, 15-25.

Boykin, A.W. & Harackiewicz, J. (1981). Epistemic curiosity and incidental recognition in relation to degree of uncertainty: Some general trends and intersubject differences. *British Journal of Psychology, 72*, 65-72.

Boyle, G.J. (1983). Critical review of state-trait curiosity test development. *Motivation and Emotion, 7*(4), 377-397.

Boyle, G.J. (1989). Breadth-Depth or State-Trait curiosity? A factor analysis of state-trait curiosity and state anxiety scales. *Personality and Individual Differences, 10*(2), 175-183.

Bruner, J.S. (1966). *Toward a theory of instruction*. Massachusetts: The Belknap Press of Harvard University Press.

Byman, R. (1993). From two-dimensional to three-dimensional curiosity: A reanalysis of depth-breadth-factor model. *Australian Journal of Psychology, 45*(3), 155-160.

Cacioppo, J.T. & Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116-131.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety.* San Franciso: Jossey-Bass.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Row.

Day, H.I. (1968). The role of specific curiosity in school achievement. *Journal of Educational Psychology, 59*(1), 37 - 43.

Day, H.I. (1971). The measurement of specific curiosity. In H.I. Day, D.E. Berlyne & D.E. Hunt (Eds.), *Intrinsic motivation: A new direction in education* (pp. 99-112). Toronto: Holt, Rinehart and Winston.

Day, H.I. (1982). Curiosity and the interested explorer. *Performance and instruction, 21*(4), 19-22.

Day, H.I. & Berlyne, D.E. (1971). Intrinsic motivation. In G.S. Lesser (Ed.). *Psychology and educational practice* (pp. 294-335). Glenview, Illinois: Scott, Foresman and Company.

Day, H.I. & Maynes, F. (1972). The teacher and curiosity: A third way. In N. Byrne & J. Quarter (Eds). *Must school fail? The growing debate in Canadian education.* Toronto: McCelland and Stewart Limited.

Deci, E.L. (1975). *Intrinsic motivation.* New York: Plenum Press.

Deci, E.L. (1992). The relation of interest to the motivation of behavior: a self-determination theory perspective. In Renninger, K.A., Hidi, S. & Krapp, A. (Eds.) *The role of interest in learning and development* (pp. 43-70). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Deci, E.L. & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Press.

Deci, E.L., Vallerand, R.J., Pelletier, L.G., & Ryan, R.M. (1991). Motivation in education: The self-determination perspective. *The Educational Psychologist, 26*, 325-346.

Dewey, J. (1913). *Interest and effort in education.* New York: Houghton Mifflin.

Fowler, H. (1965). *Curiosity and exploratory behavior.* New York: The Macmillan Company.

Gottfried, A.E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology, 77*(6), 631-645.

Gottfried, A.E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology, 82*(3), 525-538.

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in classroom: Motivational and informational components. *Developmental Psychology, 17*(3), 300-312.

Harter, S. & Jackson, B.K. (1992). Trait vs. nontrait conceptualizations of intrinsic/extrinsic motivational orientation. *Motivation and Emotion, 16*(3), 209-230.

Harty, H., Samule, K.V., & Beall, D. (1986). Exploring relationship among four science teaching-learning affective attributes of sixth grade students. *Journal of Research in Science Teaching, 23*(1), 51-60.

Haywood, H.C. (1971). Individual differences in motivational orientation: A trait approach. In H.I. Day, D.E. Berlyne & D.E. Hunt (Eds.) *Intrinsic motivation: A new direction in education* (pp. 113-127). Toronto: Holt, Rinehart and Winston.

Hebb, D.O. (1955). Drives and the C.N.S. *Psychological Review*, 62, 243-245.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60*(4), 549-571.

Hidi, S. & Anderson, V. (1992). Situational interest and its impact on reading and expository writing. In Renninger, K.A., Hidi, S. & Krapp, A. (Eds.) *The role of interest in learning and development* (215-238). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Hunt, J.M. (1971). Toward a history of intrinsic motivation. In H.I. Day, D.E. Berlyne, & D.E. Hunt (Eds.), *Intrinsic motivation: A new direction in education* (pp. 186-196). Toronto: Holt, Rinehart & Winston.

Hutt, C. (1970). Specific and diversive exploration. In H.W. Reese & L.P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 5, pp. 120-180). New York: Academic Press.

Hutt, C. (1981). Toward a taxonomy and conceptual model of play. In H.I. Day (Ed.), *Advances in intrinsic motivation and aesthetics* (pp. 251-298). New York: Plenum Press.

Iran-Nejad, A. (1987). Cognitive and affective causes of interest and linking. *Journal of Educational Psychology, 79*(2), 120-130.

Iran-Nejad, A., McKeachie, W.J. & Berliner, D.C. (1990). The multisource nature of learning: An introduction. *Review of Educational Research*, 60(4), pp. 509-515.

Izard, C.E. (1991). *The psychology of emotions*. New York: Plenum Press.

Kansanen, P. (1986). From short-term motivation to purposiveness. In T. Kroksmark & F. Marton (Eds.), *Individernas förutsättningar för utbilding* *och utbildningens effekter på individerna* (pp. 137-146). Institutionen för pedagogik Göteborgs universitet, 1986:02.

Keller, H., Schneider, K. & Henderson, B. (1994). Preface: The study of exploration. In H. Keller, K., Schneider & B., Henderson (Eds.), *Curiosity and exploration* (pp. 1-13) . Heidelberg: Springer-Verlag.

Keller, J.A. (1987). Motivational aspects of exploratory behavior. In D. Görlitz & J.F. Wohlwill (Eds.), *Curiosity, imagination, and play* (pp. 24-42). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Kirkland, J. 1976a. Interest: Phoenix in psychology. *Bull. Br. psychol. Soc 29*, 33-41.

Kirkland, J. 1976b. Epistemic curiosity and cartoon preference. *Psychological Reports 38*, 354.

Krapp, A.. (1994). Interest and curiosity. The role of interest in a theory of exploratory action. In H. Keller, K. Schneider & B. Henderson (Eds.), *Curiosity and exploration* (pp. 79-99). Heidelberg: Springer-Verlag.

Krapp, A., Hidi, S. & Renninger, K.A. (1992). Interest, learning and development. In Renninger, K.A., Hidi, S. & Krapp, A. (Eds.) *The role of interest in learning and development* (pp. 3-25). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Kreitler, S., Kreitler, H. & Zigler, E. (1975). The nature of curiosity in children. *Journal of School Psychology, 13*(3), 185-200.

Langevin, R. (1971). Is curiosity a unitary construct? *Canadian Journal of Psychology, 25*, 360-374.

Leuba, C. (1955). Toward some integration of learning theories: The concept of optimal stimulation. *Psychological Reports, 1*, 27-33.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75-98.

Madsen, K.B. (1974). *Modern theories of motivation*. Copenhagen: Munksgaard.

Maw, W. H. (1971). Differences in the personality of children differing in curiosity. In H. I. Day, D. E. Berlyne, & D. E. Hunt (Eds.). *Intrinsic motivation: A new direction in education* (pp. 91-98). Toronto: Holt, Rinehart and Winston of Canada, Limited.

Maw, W.H. & Maw, E.W. (1965). Personal and social variables differentiating children with high and low curiosity. Cooperative Research Project No. 1511, University of Delaware.

Miller, M.V. (1987). Curiosity and its vicissitudes. *The Gestalt Journal*, 10(1), pp. 18-32.

Naylor, F.D. (1981). A state-trait curiosity inventory. *Australian Psychologist, 16*(2), 172-183.

Nunnally, J.C. (1981). Explorations of exploration. In H.J. Day (Ed.), *Advances in intrinsic motivation and aesthetics* (pp. 341-364). New York: Plenum.

Olson, K. & Camp, C. (1984). Factor analysis of curiosity measures in adults. *Psychological Reports, 54*, 491-497.

Pearson, P. (1970). Relationships between global and specified measures of novelty seeking. *Journal of Consulting Psychology*, 11, 199-204.

Penny, R.K. (1965). Reactive curiosity and manifest anxiety. *Child Development, 36*, 697-702.

Piaget, J. (1981). Intelligence and affectivity: Their relationship during child development. Palo Alto, California: Annual Reviews Inc.

Rigby, C.S., Deci, E.L., Patrick, B.C., & Ryan, R.M. (1992). Beyond the intrinsic-extrinsic dichotomy: Self-determination in motivation and learning. *Motivation and Emotion, 16*, 165-185.

Rubenstein, A. S. (1986). *An item-level analysis of questionnaire-type measures of intellectual curiosity.* Unpublished doctoral dissertation, Cornell University.

Schafer, R. (1968). *Aspects of internalization.* New York: International University Press.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, 26, 299-323.

Schiefele, U. (1992). Topic interest and levels of text comprehension. In Renninger, K.A., Hidi, S. & Krapp, A. (Eds.) *The role of interest in learning and development* (pp. 151-182). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Schiefele, U., Krapp, A. & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K.A Renninger, S. Hidi & A. Krapp, (Eds.) *The role of interest in learning and development* (pp. 183-212). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Schneider, K. & Unzner, L. (1994). Preschoolers' exploratory behavior: The influence of the social and physical context. In H. Keller, K. Schneider & B. Henderson (Eds.), *Curiosity and exploration* (pp. 177-197). Heidelberg: Springer-Verlag.

Scriven, M. (1988). Philosophical inquiry methods in education. In R.M. Jaeger (Ed.), *Complementary methods for research in education* (pp. 131-183). Washington, DC: AERA.

Snow, R.E. & Jackson, D.N. (1994). Individual differences in conation: Selected constructs and measures. . In H. F. O'Neil, Jr. & M. Drillings (Eds.), *Motivation: Theory and research* (pp. 71-99). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Spielberger, C.D. & Starr, L.M. (1994). Curiosity and exploratory behavior. In H. F. O'Neil, Jr. & M. Drillings (Eds.), *Motivation: Theory and research* (pp. 221-243). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64*(1), 37-54.

Tolentino, E., Curry, L. & Leak, G. 1990. Further validation of the short form of the need for cognition scale. *Psychological Reports 66*, 321-322.

Trudewind, C. & Schneider, K. (1994). Interindividual differences in the development of exploratory behavior. Methodological considerations. In H. Keller, K. Schneider & B. Henderson (Eds.), *Curiosity and exploration* (pp. 154-176). Heidelberg: Springer-Verlag.

Vallerand, R.J., Pelletier, L.G., Blais, M.R., Briére, N.M., Senécal, C. & Valliéres, E.F. (1992). The academic motivation scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement*, 52, 1003-1017.

Valsiner, J. (1992). Interest: A metatheoretical perspective. In Renninger, K.A., Hidi, S. & Krapp, A. (Eds.) *The role of interest in learning and development* (pp. 27-41). Hillsdale, New Jersey: Lawrence Erlbaum Associates, inc.

Voss, H.-G. (1987). Exploration and play: Research and perspectives in developmental psychology. In D. Görlitz & J.F. Wohlwill (Eds.), *Curiosity, imagination, and play* (pp. 43-58). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Voss, H.-G. & Keller, H. (1983). *Curiosity and exploration. Theories and Results.* New York: Academic Press, Inc.

Weiner, B. (1990). History of motivational research in education. *Journal of Educational Psychology, 82*(4), 616-622.

Weiner, B. (1992). Motivation. In M. C. Alkin (Ed.) *Encyclopedia of Educational Research* (pp. 860-865). New York: Macmillan Publishing Company.

White, R.W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*, 297-333.

Wittgenstein, L. (1980). *Sininen ja ruskea kirja* [The blue and brown books: preliminary studies for the "Philosopical investications"]. Juva: WSOY.

Wittgenstein, L. (1981). *Filosofisia tutkimuksia* [Philosophical investigations]. Juva: WSOY.

Wohlwill, J.F. (1981). A conceptual analysis of exploratory behavior: The "specific-diversive" distinction revisited. In H.J. Day (Ed.), *Advances in intrinsic motivation and aesthetics* (pp. 341-364). New York: Plenum.

Wohlwill, J.F. (1987). Varieties of exploratory activity in early childhood. In D. Görlitz & J.F. Wohlwill (Eds.), *Curiosity, imagination, and play* (pp. 59-77). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Woodworth, R.S. (1918). *Dynamic psychology*. New York: Colombia University Press.

Woodworth, R.S. (1958). *Dynamics of behavior*. New York: Holt, Rinehart and Winston.

von Wright, J. (1986). Kognitiivisesta oppimiskäsityksestä. *Psykologia*, 21, 83-88.

Zuckerman, M. (1971). Dimensions of sensation seeking. *Journal of Consulting and Clinical Psychology*, 36, 45-52.

Zuckerman, M. (1979). *Sensation seeking: Beyond the optimum level of arousal*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Zuckerman, M. (1984). Sensation seeking: A comparative approach to a human trait. *The Behavioral and Brain Sciences*, 7, 413-434, 453-471.

Zuckerman, M. (1987). A critical look at three arousal constructs in personality theories: Optimal levels of arousal, strength of the nervous system, and sensitivities to signals of reward and punishment. In J. Strelau & H.J. Eysenck (Eds.), *Personality dimensions and arousal*. (pp. 218-232). New York: Plenum Press.

The Development of Initial Teacher Education: Insights from Research on Learning to Teach

James Calderhead

In several countries, radical changes are being proposed for teacher education. In the UK, USA and Australia, there is a widespread belief that teacher education should be more "practical" and school-based and that teaching effectiveness would be improved if more teachers themselves were involved directly in the training process. Interestingly, many of the ideas about teaching and teacher education, and also about the processes of educational innovation and implementation, that are implicit in current proposals for reform, contrast sharply with those that emerge from the research literature. The aim of this paper is to examine those contrasts and to consider the contribution of research to the advancement of quality in teacher education.

Recent policy documentation in the UK, in particular DfE Circular 9/92 with its list of teacher competences, reflects quite specific conceptions of teaching and teacher education. Teaching appears to be thought of principally in terms of the possession of subject matter knowledge and its transmission to children. Particular importance is attached to the subject expertise of teachers and how this is applied in instruction and assessment, while teacher education is seen largely in terms of ensuring student teachers' subject mastery and their acquisition of the classroom skills of planning, management, presentation and assessment. Furthermore, the process of reform of teacher education seems to be viewed as one of merely changing accreditation requirements and of introducing a market economy to the supply and demand for teachers. By developing alternative routes into teaching and financial incentives for schools themselves to provide training, it is assumed that the resulting competition amongst providers will lead to improved quality. Such views of teaching and teacher education should not surprise us. After all, many student teachers start out with similar views, believing that teaching is largely a matter of recounting facts to children, and that learning is a matter of memorising. Consequently, learning to teach is viewed in terms of knowing one's subject and picking up some practical teaching tips. Such views have been nourished by the student teachers' observations of their own teachers over several years. These are common, stereotyped conceptions that frequently find expression in everyday parlance, but which are difficult to sustain when examining, in detail, what is actually taught and learned in schools. They suggest that teaching and learning are more transmission-oriented, subject focussed, decontextualised, predictable and static than they really seem to be, and such views contrast sharply with those that emerge from detailed observational or ethnographic studies of classrooms. Walter Doyle (1986), for example, based on his own ethnographic work in schools, suggests that teaching has 6 major distinguishing characteristics: *multidimensionality, simultaneity, immediacy, unpredictability, publicness* and *history*. Classrooms, in other words, are busy places, with lots of activities going on, many different agendas being acted out and, as a result, choices for the teacher are never simple. There are costs and benefits associated with even everyday, relatively trivial decisions. Waiting a few extra moments for a child to answer a question, for example, may encourage that child to think more deeply and to think through a problem for himself, but it might be at the cost of losing the attention of others in the group or class. Several activities may also be occurring at the same time and teachers may find themselves engaged in more than one task simultaneously, asking a question that is pertinent to a particular stage in a lesson, and which signals a key issue, for example, but asking it in a tone of voice and addressing it to a particular pupil that, at the same time, signals that the teacher has noted his inattention. Teachers may need to respond immediately to situations - they cannot wait, or go away and think about it and come back, they are in demanding situations and have to respond or intervene with little advance notice. And teachers can never be sure exactly what is going to happen next or what effects their actions might have. One of the mysteries of children's learning is that whilst they might have understood the second law of thermodynamics yesterday, this is no guarantee that they will understand it today. And as well as being unpredictable, classrooms are public places and teachers are on view - how they react to events can affect how the whole class later responds. The teacher who reprimands an innocent pupil or ignores the violation of a class rule, for example, is providing signals to the rest of the class that may well influence how they think of the teacher and how they relate to him/her in the future. Finally, over the weeks of interaction between a teacher and a class, the class develops its own norms, its own ways of working, its own expectations - a history that shapes the everyday working practices of teachers and pupils.

The view of teaching that Walter Doyle paints for us is one of a dynamic, complex process that is extremely demanding of teachers - behaviourally, cognitively and affectively. Teachers have *to act*, they need strategies and routines to respond to the variety of situations they face. They also need *to know* when to employ those routines, to be able to analyse situations and identify when to act, and also to be able to evaluate past practice so that they can note where things have gone wrong and how they might proceed. Teachers also need to be able *to empathise* with pupils and to establish relationships with children, colleagues and parents to develop the climate and situation in which teaching and learning can take place.

An indication of some of this complexity of teaching, and of the differences in the ability to cope with this on the part of student and experienced teachers, was evident in some research pursued by the author (Calderhead, 1981). In this study of primary school teachers' thinking and decision-making, each teacher was presented with a list of classroom critical incidents covering a wide range of typical managerial, instructional and social situations. After each incident, they were asked what more they needed to know about the situation in order to decide how to deal with it and what they would do.

Experienced teachers would generally recall a series of typical incidents rather like the one presented to them and suggest a few strategies for coping. Student teachers, on the other hand, usually gave much simpler responses. For example, one of the incidents was 'The class is working quietly when a group of children start talking amongst themselves'. Experienced teachers frequently recounted six or more typical incidents of this kind - the class clown playing up, the child having a difficulty and asking around for help rather than approaching the teacher, a child not having listened to instructions and having to ask others what to do, a distraction like a wasp flying in through the window, or another teacher coming into the room, etc.

In contrast, one student teacher responded to the same situation, saying: "I'd wait until the noise reached an intolerable level then I'd tell them to shut up" - and encouragement to expand on that response resulted in little elaboration.

The experienced teachers had more sophisticated perceptions of the critical incidents, and perceived the incidents in terms of an example of a possible range of typical situations, and this range could be very readily brought to mind.

Interestingly, the experienced teachers did not provide anything like the same number of action solutions as they did variants of the problem. One interpretation of this is that teachers' knowledge is not held in the form of action recipes for typical situations - "this is what you do in x type situation." Teachers use their memory of classroom incidents as a way of thinking about the particular problems they encounter. Their response to any particular situation may depend as much on their assessment of the children and context involved as the typifications they call to mind. When a group of children start making a lot of noise, for example, how the teacher responds may depend on a variety of factors, including what has happened immediately before, the type of lesson it is, the children involved, whether the children are working individually or in groups, and what seems likely to happen next, or how close it is to the end of the school day, etc. This may well indicate something of the nature of teachers' knowledge and the way it is used to understand and respond to classroom situations.

Given the complexity of teaching, it is not surprising that learning to teach, for many students, turns out to be difficult and troublesome. That complexity, however, is often underestimated or misunderstood by the student teachers themselves. Commonly they perceive learning to teach in terms of acquiring a repertoire of recipes, and in consequence they accumulate and reproduce observed strategies, sometimes without much recognition of the context in which they are being used.

In our efforts to support students learning to teach it would obviously be useful to have some global understanding of what is involved - what learning occurs, what difficulties students typically encounter, and how we help students to overcome them.

Unfortunately there is no one comprehensive theory of learning to teach. The whole area is perhaps too complex for such a synthesis currently to be made. There are, however, a number of different theoretical models that help us to conceptualise particular aspects of the processes involved, and shed light on some of its complexities.

1. Socialisation into the Professional Culture

Several researchers have emphasised the socialising processes in professional development. Teaching is portrayed as a complex task that occurs in a powerful material and ideological context. The organisation and physical resources of schools and, perhaps more significantly, the beliefs that are not only held and valued within the institution but have become embedded within its many taken-for-granted practices, inevitably exert a powerful influence upon the new teacher, sometimes referred to as the "wash out effect" since they may dominate and over-rule the practices learned in college. Lacey (1977) describes the strategies adopted by new teachers in school to negotiate their way amongst these pressures, changing to fit in with some of the school values and practices, managing to influence others, and going along with some for strategic reasons in order, for example, to gain acceptance within the school even though not essentially believing in them. Zeichner, Tabachnick and Densmore (1987) propose that the socialisation process is in fact more complex because schools consist of several cultures and the task of the new teacher is to weave their way amongst them often identifying with one like-minded individual or group of teachers from whom they can seek support. Zeichner and Gore (1990) suggest that there are several models of teacher socialisation, which place differing emphases on the influence of the institutional context and the individual teacher's potential to change existing values and practices. Socialisation perspectives on professional development, however, have succeeded in highlighting the complex interactions that occur between an individual's own values, beliefs and practices and those of the school, and also the importance of the individual's capacity to negotiate and manoeuvre within a powerful ideological context. This may well raise questions about the importance of 'political skills' amongst student teachers, and how students might be better prepared to develop and defend their own practice within an institutional context.

2. The development of knowledge and skills

This is perhaps the most often cited perspective on learning to teach which emphasises the knowledge and skills that contribute to classroom practice. In the 1960s and 70s, this was defined in terms of classroom behaviours. Several attempts were made to develop microteaching or skill development programmes, relying particularly upon the large amount of process-product research that attempted to link teaching behaviours to children's learning outcomes. More recently teaching skill has been conceptualised cognitively as well as behaviourally, in terms of ways of understanding practice and the actions to which such understandings lead. Several studies comparing experienced and novice teachers have demonstrated how the experienced teacher has a much more sophisticated understanding of their practice. The experienced teacher appears to have access to a wide range of knowledge that can be readily accessed when dealing with classroom situations and which can help in interpreting and responding to them. The novice teacher, on the other hand, makes simpler, commonsense interpretations of classroom events, and is less able to anticipate possibilities and act accordingly (e.g. Berliner, 1987).

The knowledge that teachers draw upon in planning their work, teaching in the classroom and later evaluating their work is highly varied, including knowledge of children, teaching strategies, the curriculum, school rules, the availability of materials, etc. Teachers' knowledge also takes different forms, some of it is clearly propositional and is relatively easily articulated, other aspects of teachers' knowledge are embedded within action and are more easily demonstrated than talked about. One aspect of teachers' knowledge that has received particular attention in recent years has been teachers' knowledge of subject matter. Largely stimulated by the work of Shulman (Shulman, 1986; Wilson, Shulman & Richert, 1987), several studies have investigated changes in student teachers' understanding of the subject they teach. These have suggested that a large amount of time is spent in the first few years of teaching re-learning the subject matter for the purposes of teaching it. Shulman suggests that in order to teach a subject one needs both a breadth and depth of knowledge, a rich factual knowledge base with many interconnections which represents a much more thorough understanding than that which is achieved purely as a learner. In addition, teachers develop knowledge which enables them to facilitate understanding in others. Shulman refers to this as *pedagogical content knowledge* - knowledge of examples, anecdotes, experiments, and difficulties that are commonly experienced that help teachers to communicate about the subject matter. Since children's own backgrounds vary considerably and they approach a subject with particular understandings of their own, teachers, it is argued, need a wide repertoire of pedagogical content knowledge to cater for children's individual differences. The analogy that 'works' for one child, for example, may be completely meaningless to another.

One example of pedagogical content knowledge that Shulman and his colleagues present is of a novice teacher faced with the task of teaching Julius Caesar to a class of 14 year olds. Believing that the children will not be enthusiastic about the play, he uses his knowledge both of the play as a drama about moral conflict and of the children and their interests to construct a task where he asks the children to imagine they are second in command on the Starship Enterprise and that Captain Kirk, their best friend, is becoming big-headed and is beginning to take decisions that risk the lives of his crew - what would they do? This situation is used to lead into a discussion on loyalty and morality. The teacher uses his knowledge of the subject, and of the children and their interests, to introduce the play in a way that captures their attention, engages them in considering the theme of the play, and enables them to empathise with the characters, setting the scene for more in-depth study. In the process, the teacher also adds to his own pedagogical content knowledge. Developing this knowledge not only requires an understanding of the subject, but an understanding of children, their abilities and interests and how they tend to respond to different situations, a knowledge of teaching strategies and how different types of classroom activity might be managed. Developing pedagogical content knowledge, in fact, seems to require the orchestration of a wide variety of knowledge about teaching.

Studies of novice and experienced teachers suggest that there is an enormous diversity of knowledge that the more experienced teacher possesses - not only about subject matter, but about children, teaching and the classroom context - that enables teachers to make sense of classrooms and to monitor and shape their classroom routines and behaviours.

3. The moral dimension of teaching

Teaching as well as being a practical and intellectual activity is also a moral endeavour. Teaching involves caring for young people, considering the interests of children, preparing them to be part of a future society, and influencing the way in which they relate to each other and live. The ethic of caring has been claimed to be a central facet of teaching (Noddings, 1986; Valli, 1990), often valued by teachers, parents and children, but frequently unacknowledged in discussions of professional development. In addition, important moral issues concerning equality within our society are implicit in the way in which schools and teaching are organised and the ways in which academic rewards, in the form of qualifications, are distributed, and this can raise several moral dilemmas for teachers in their practice.

Olson (1992) suggests that in many countries recent concerns with educational standards and instrumental attempts to manipulate the processes of teaching and learning towards certain areas of scholastic attainment have resulted in the moral dimension of teaching becoming obscured. He argues that implicit in all teaching acts are certain value judgments and that in education today these may well be at odds with the basic values to which teachers, and teaching as a profession, might wish to adhere. Part of a teacher's professional development, he suggests, is to enquire into what those values are, to regain control of those values and to seek to promote a closer level of agreement between values and practices.

4. The personal dimension of teaching

Several studies have recently emphasised the close relationship between the personal and the professional in teachers' work, and the need to consider personal development in any consideration of teachers' professional development. Several different aspects of the personal dimension have been emphasised in the literature. First of all, it has been found that teachers approach teaching with various ideas and images of what teachers' work is like based on their own individual past experiences, including previous work experience, experiences as a parent or childhood experiences of school (Johnston, 1992; Calderhead and Robson, 1990). This is often a metaphoric way of thinking about teaching (Russell et al, 1988) in which past experiences are used as models for reasoning about the nature of teaching and what teachers do. In case studies of experienced teachers, Elbaz (1983) and Clandinin (1986) have used the term 'personal practical knowledge' to refer to the working knowledge that teachers have that is imbued with teachers' own lived experience.

Secondly, teachers' personalities themselves are an important aspect of their work as teachers (Clark, 1988). Teachers are to some extent performers in the classroom, required to establish working relationships with their pupils, to command their attention and respect and to ensure the smooth running of their classes. Such tasks intrinsically involve teachers' personalities, and part of teachers' professional development requires teachers to engage in self-learning, becoming aware of their own personal qualities and how other people respond to them, so that they can take greater control in their interactions with others.

5. The reflective dimension

Reflection has recently become a popular term in the context of teachers' professional development. Many teacher education courses, both preservice and inservice, claim to be based upon a reflective practitioner model. The enthusiasm for reflective practice may be partly accounted for in terms of the

BEST COPY AVAILABLE

current attractiveness of many of the principles that have come to be associated with it: helping teachers to analyse, discuss, evaluate and change their own practice; heightening teachers' awareness of the contexts in which they work; enabling teachers to appreciate the moral and ethical issues implicit in their practice; to empower teachers to take greater control over their own professional growth and to influence future directions in education. At a time of uncertain educational change and when, in some countries, there are strong pressures to view teachers as technicians and implementers of a curriculum devised by others, notions of teachers as reflective practitioners - educators with a purpose, able to take control over their own futures - has an obvious appeal. What reflective practice actually looks like, however, is less clear. Various notions of reflection have been used to support a diversity of practices.

One common conception of reflection in teaching refers to the processes of monitoring and adapting one's behaviour in context. Schon (1983, 1987) referred to this as *reflection-in-action*. Because teaching is complex and unpredictable, teachers cannot rely entirely on routine ways of coping with situations. Teaching therefore involves a process of acting, reflecting on the effects of one's actions and constantly adapting one's behaviour to the situation and purposes at hand.

Another conception of reflection places emphasis on after-the-event evaluation. Teachers, after a lesson or after the day is over, may reflect back on particular events, analysing where difficulties arose, considering how they might be surmounted and deciding on the future directions their teaching might take. In Schon's terminology, this would be *reflection-on-action*.

Several researchers (e.g. Tom, 1985; Carr & Kemmis, 1986) have also pointed towards a more deliberative form of reflection that teachers engage in - reflection that is more searching, philosophical and critical. These are occasions when teachers think more about the purposes of education, their own personal and professional goals and the value of their own practice.

Recent writings on reflective teaching have considered what the cognitive, affective and behavioural components of reflection might be: what are the skills, knowledge bases, attitudes and predispositions that make reflection possible, and how might these be facilitated amongst teachers at different stages in their career? Do student teachers need to be taught how to reflect and is this different from the support needed by more experienced teachers in their reflection? Do student teachers need to have a basic mastery of teaching before reflection on practice is actually possible? Researchers and teacher educators have also focused on the content of reflection: is all content an equally appropriate subject for reflection, and how does one prioritise the subject matter for reflection? (e.g. see Calderhead and Gates, 1993)

Certainly, enthusiasm for reflective teaching has led to much experimentation in teacher education (see Clift, Houston & Pugach, 1990; Valli, 1992) and the development of a wide range of techniques and approaches to encourage and foster reflection in both preservice and inservice courses. These have included techniques such as journal writing, action research, and the use of research evidence and empirically derived theory to provide alternative conceptual frameworks for the analysis of practice, as well as the development of certain principles of training and defined roles of the trainers that relate to the overall 'reflective' philosophy and organisation of a course.

The implementation of reflective teacher education has not been without its difficulties, however. Creating a course that helps student teachers to become more analytical about their practice and to take charge of their own professional development is a task with a number of inherent dilemmas. How do teacher educators reconcile their traditional role as a gatekeeper to the profession with that of mentor and facilitator of reflection? The goals of reflective teaching are extremely ambitious - what is reasonable to achieve in preservice education and what can only be achieved in the much longer term? How does a teacher education institution foster reflection when in schools much greater importance is attached to immediate, spontaneous action than to reflection and evaluation? Does reflective teaching require a particular supportive, collaborative ethos in school in order for the efforts of teacher educators to be effective? The development of reflective practice in preservice and inservice courses has not proved easy and there are many questions posed by current research and development efforts.

From research on learning to teach, we can conclude that teaching, and the processes of learning to teach, are highly complex. Each of the dimensions discussed above identifies an important set of variables relating to professional development, but each only gives a partial picture of the total process. Each focuses on a particular aspect of teachers' work, though clearly all of these dimensions of teachers' practice are closely inter-related. Learning to teach involves the development of technical skills, as well as an appreciation of moral issues involved in education, an ability to negotiate and develop one's practice within the culture of the school, and an ability to reflect and evaluate both in and on one's actions. There are different forms of learning required and this presents difficulties for the beginning teacher. First of all, the learning required of them is often quite different from the learning in which they have previously been engaged in higher education: learning to become a teacher is different from learning mathematics or learning history, for example. Secondly, learning to be a teacher demands multiple forms of learning. Learning to teach the concept of ratio is different from learning to present oneself as a teacher in the classroom, or learning to relate to reluctant learners, or learning how to plan the curriculum, or how to work with one's colleagues or how to cope with one's own anxieties (see Calderhead, 1991). Not only does learning to teach involve different forms of learning, but since student teachers start out with many different abilities, types of expertise and background experience, their routes in the process of learning are inevitably quite varied.

There is an important research agenda here for teacher education. We need to explore those processes of learning to teach more fully, to understand what it means to become a teacher and to know how to support teachers' development.

There is also an important political agenda, however - as teacher educators, we may not in the past have been very effective in communicating to others what teaching and teacher education are about, and we may well have been less than fully cognisant of the need to do this. As a result, we lack the vocabulary and the conceptual frameworks to demonstrate to others that teaching is a complex and challenging task, and that teacher education deserves to be taken seriously. There is a need to be able to convince policy-makers and others of what teacher education involves, to counter the simple rhetoric that abounds, and to gain recognition for the demanding work of teacher education as we know it.

References

Berliner, D.C. (1987) Ways of thinking about students and classrooms by more and less experienced teachers. In J. Calderhead (Ed.) *Exploring Teachers' Thinking*. London: Cassell.

Calderhead, J. (1981) A psychological approach to research on teachers' classroom decision-making. *British Educational Research Journal*, 7 (1), 51-58.

Calderhead, J. (1991) The nature and growth of knowledge in student teaching. *Teaching and Teacher Education*, 7 (5/6), 531-535.

Calderhead, J. & Gates, P. (1993) *Conceptualizing Reflection in Teacher Development*. London: Falmer Press.

Calderhead, J. & Robson, M. (1990) Images of teaching: student teachers' early conceptions of classroom practice. *Teaching and Teacher Education*, 7 (1), 1-8.

Carr, W. & Kemmis, S. (1986) *Becoming Critical: education, knowledge and action research*. London: Falmer Press.

Clandinin, D. J. (1986) *Classroom Practice: teacher images in action*. London: Falmer Press.

Clark, C.M. (1988) Asking the right questions about teacher preparation: contributions of research on teacher thinking. *Educational Researcher*, 17 (2), 5-12.

Clift, R., Houston, W.R. & Pugach, M. (1990) *Encouraging Reflective Practice in Education*. New York: Teachers' College Press.

Doyle, W. (1986) Classroom organization and management. In M.C. Wittrock (Ed.) *Handbook of Research on Teaching, third edition*. New York: Macmillan.

Elbaz, F. (1983) *Teacher Thinking: a study of practical knowledge*. London: Croom Helm.

Johnston, S. (1992) Images: a way of understanding the practical knowledge of student teachers. *Teaching and Teacher Education*, 8 (2), 123-136.

Lacey, C. (1977) *The Socialisation of Teachers*. London: Methuen.

Noddings, N. (1986) Fidelity in teaching, teacher education and research on teaching. *Harvard Educational Review*, 56 (4), 496-510.

Olson, J. (1992) *Understanding Teaching: beyond expertise*. Milton Keynes: Open University Press.

Russell, T., Munby, H., Johnston, P. & Spafford, C. (1988) Learning the professional knowledge of teaching: metaphors, puzzles and the theory-practice relationship. In P. Grimmett & G.L. Erickson (Eds.) *Reflection in Teacher Education*. New York: Teachers' College Press.

Schön, D.A. (1983) *The Reflective Practitioner*. London: Temple Smith.

Schön, D.A. (1987) *Educating the Reflective Practitioner*. New Jersey: Jossey-Bass.

Shulman, L.S. (1986) Those who understand: knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.

Tom, A. (1985) Inquiring into inquiry-oriented teacher education. *Journal of Teacher Education*, 36 (5), 35-44.

Valli, L. (1992) *Reflective Teacher Education: cases and critiques*. New York: SUNY Press.

Wilson, S.M., Shulman, L.S. & Richert, A. (1987) '150 different ways' of knowing: representations of knowledge in teaching. In J. Calderhead (Ed.) *Exploring Teachers' Thinking*. London: Cassell.

Zeichner, K.M. & Gore, J. M. (1990) Teacher socialisation. In W.R. Houston (Ed.) *Handbook of Research on Teacher Education*. New York: Macmillan.

Zeichner, K.M., Tabachnick, B.R. & Densmore, K. (1987) Individual, institutional and cultural influences on the development of teachers' craft knowledge. In J. Calderhead (Ed.) *Exploring Teachers' Thinking*. London: Cassell.

The Role of School Practice in Teacher Education

Juhani Hytönen

1. Background

Teacher education programmes were renewed in Finland in 1979. The programmes for primary teacher (classroom teacher) education were changed to a four-and-a-half-year programme in the university. The new programme consists of more theoretical studies - mostly educational theory. An outcome from this is that a student teacher must write a thesis in education and complete an M.A. The programme has about the same amount of practice in teaching as the earlier one. Qualitatively the present practice, however, differs from the old school practice.

The main task of theoretical studies in a teacher education programme is to help student teachers to understand their teaching experiences and the problems they will face in the classroom. In the previous training programme the amount of theoretical studies was smaller and the level of theoretical studies was lower than today. Student teachers were mainly given practical, ready-made and tried-out models during their theoretical preparation. Accordingly, in teaching situations supervising teachers acted according to certain prototypes. The training solution was a concrete and safe one. It also made student teachers' teaching behaviour uniform. The new teacher education programme can at its best give student teachers examples with reasons of how a teacher can act in a teaching situation. The programme can also support student teachers when they begin to find their personal ways to solve the kinds of problems which teachers face in their work.

2. The theoretical framework of school practice

The new programme of classroom teacher education in Helsinki was originally based on the ideology of humanistic psychology (Hytönen 1982, 21-22). A student teacher's personality ("self") and confidence in her/himself as a teacher were key elements in the supervision of practice teaching. The general aim of the primary teacher education programme could be reduced to the concept of "the reflective teacher". The goal was to develop an innovator-teacher who could analyse teaching situations with the help of scientific thinking.

Gradually the education goal based on humanistic psychology was found to be too individually centred, and an innovator-teacher as an ideal teacher to be too narrow in scope (Hytönen 1989). Liston and Zeichner (1991, 38) describe appropriately our conception when they write, *"lately the sense within teacher education seems to be that as long as teachers 'reflect' on their actions and purposes, everything is all right. When this is the case, calls for further reflection become groundless - that is, they lack a substantial basis for discerning what will count as good reasons for educational actions. Since teachers are (and we think should be) responsible for key educational decisions and actions within the classroom, it is important that prospective teachers begin to consider what will count as a good reason for effective educational action."*

We educate teachers in a rapidly changing society and world. Teachers must be able to analyse and develop school as a part of society and even society as part of the world, not only to transfer the cultural inheritance. Student teachers should already during their training period understand that there are factors in school and society which need to be reformed. Hence the practice teaching periods must also contain a critical view of the teacher's profession. Teacher education (and school practice) needs a wider base for a theory where the social, philosophical and historical aspects of education are seen together with the didactic view. The current (and future) school needs teachers who actively participate in the development of a more democratic society. (Hytönen 1989.)

Recently we have searched for theoretical ideas and concepts, as well as methodological solutions, mainly from the social-reconstructionist tradition (i.e. critical pedagogy) in teacher education (see e.g. Giroux & McLaren 1986, Liston & Zeichner 1991 and Zeichner 1991). Critical pedagogy views schools as democratic public spheres and teachers as transformative intellectuals. McLaren (1989, 238) writes that *"viewing schools as democratic public spheres means regarding schools as sites dedicated to forms of self and social empowerment, where students have the opportunity to learn the knowledge and skills necessary to live in an authentic democracy. Instead of defining schools as extensions of the workplace or as frontline institutions in the battle for international markets and foreign competition, schools as democratic public spheres function to dignify meaningful dialogue and action and to give students the opportunity to learn the language of social responsibility. Such a language seeks to recapture the idea of democracy as a* ***social movement*** *grounded in a fundamental respect for* ***individual freedom*** *and* ***social justice****."*

BEST COPY AVAILABLE

3. The principles of integrating theory and practice

One of the main problems in teacher education is how to integrate theoretical studies with practice teaching. Because of the greater amount of theoretical studies in the new teacher education programme, we took the question of integration very seriously from the very start of our planning. The integration principles we applied in planning our programme were (1) **integration presupposes differentiation** and (2) **integration must be carried out mainly through practice towards theory**.

Although the integration of theory and practice in teachers' daily work is a continuous and a holistic process, student teachers must get an organized perception of the process of integration in a teacher education programme when learning how to teach. Broader units will be built on the mastery of such details. However before starting the differentiation, student teachers should have a general outline of a teacher's field of work, within which details will be combined together.

Peters (1977, 139-140) sees the background of the first principle very clearly: *"Logically speaking all questions of educational policy and practice are hybrid questions, a crossing of value judgments with different forms of empirical enquiry. So the first step towards answering such questions must be the breaking down of the questions into their logically distinct components. If, for instance, we want to know whether we ought to punish children, the ethical issues of principle must be distinquished from the psychological and social questions about causes and consequences. But once the various voices that speak about education, emanating from the different disciplines, have been separated out and drained, they need to join together in an orderly and coherent conversation about matters of common concern to teachers and practical administrators. In brief, we must make an end of the undifferentiated mush that is often perpetrated under the heading of educational theory ..."*.

Behind the second principle - integration must be carried out mainly through practice towards theory - we have the axiom that teacher education is an activity which is characterized by an implicit rather than explicit utilization of theory (see Hytönen 1982, 34-38). In teacher education subjective experiences - founded on practice - should be the starting point of training; theoretical studies give student teachers additional means and support when they try to organize and understand practical situations and problems they have faced. A theory of education differs from a theory of engineering.

4. The solutions of integration

In realizing the curriculum of our classroom teacher programme, we have used the following solutions in integrating theoretical studies with practice teaching.

1) Practice teaching must be started as early as possible. Student teachers have their first practice teaching period already in their first study year. The aim is to familiarize each student teacher with the work that a teacher does. The intention is also to motivate student teachers. Starting practice teaching so early works under the presumption that student teachers have a sufficient amount of supervision and help in preparing their lessons. Teaching situations should be limited in time and character, too.

2) Interaction between theory and practice must be continuous. In our primary teacher education practice, teaching periods have been divided into periods throughout the four-and-a-half-year programme. The first academic year includes a two-week period, the third year an eight-week period, the fourth year a five-week period and the fifth year a five-week period.

We have divided our practice teaching into four periods because we assume that student teachers - after having got more theoretical studies - will in every phase view their teaching experiences from a changed perspective. Student teachers' visions about the teacher's profession and about themselves as teachers become gradually more accurate. If we had only one or even a few longer practice periods, the risk of practice teaching becoming static in form would be prevalent. Student teachers would not grow as professionals because the support of theoretical studies would be lacking. On the other hand, a long practice period creates an image of practice without theory and the integration of theory and practice becomes more difficult.

3) Each practice period has aims and a character of their own. Traditionally, practice teaching in Finnish teacher education programmes has been rather holistic in nature: a few long periods with no precise aims, "just practicing". By defining aims for each practice teaching period, we have tried to make practice teaching more effective. Both supervisors and student teachers know what to pay attention to in each practice teaching period. The different aims of each practice period give us the possibility to organize the supervisors' work, too.

The focus of the first practice period is on pupils, the teacher and the (basic) forms of interaction between them. Because theoretical studies in the social connections of schooling and in the subjects of the comprehensive school curriculum have only been initiated, such topics should be handled with careful supervision. Supervising teachers and (university) lecturers in educational theory are in charge of the supervision.

During the second practice teaching period - in the third academic year - student teachers practise how to teach subjects in the comprehensive school curriculum. Supervision goes under the leadership of supervising teachers and (university) lectures, who are specialized in the didactics of the subject in question. The second practice period is carried out in a phase when student teachers have finished their basic studies in the comprehensive school subjects.

The aim of the third practice teaching period - in the fourth academic year - is to give student teachers experiences which will help them find their personal ways of acting as a classroom teacher. Student teachers keep a journal related to their development as teachers during this practice teaching period and are given guidance in their efforts by (university) lecturers in educational theory. Student teachers are sent to practice in schools where activities are known to be of innovative or of experimental character and student teachers can choose their school according to their personal interests. Child-centredness in education and integration of the curriculum have naturally been very popular topics. However, co-operation between school and home and between school and society have been of interest during the last few years, too. Supervising teachers have done the supervising work alone.

The aim of the fourth practice teaching period - at the beginning of the fifth academic year - is for student teachers - rather autonomously, but still under supervision - to plan and carry out their teaching activities during the period. The whole area of the reflective teaching profession is within the scope of the activities. It is emphasized that reflection needs to reach beyond the particulars of student teachers' own actions and inclinations to consider larger organized contexts relevant to their future work as a primary school teacher. Student teachers are supervised by supervising teachers and (university) lecturers in educational theory.

4) Practice teaching should progress from simple elements towards complex unities. This solution of integration is implicitly seen in the solution above. While pupils, the teacher and the basic modes of interaction between them are in focus at the beginning of practice teaching, the final practice period gives a student teacher the opportunity to experience innovative teaching solutions and experiments as well as an opportunity to take her/his responsibility for educating pupils and for developing school curricula.

5) The theoretical studies which give support to a particular practice teaching period must be written out in the curriculum of the teacher education programme. The aim is to point out those theoretical courses which help student teachers to analyse what problems they have experienced during the particular practice period. For instance, theoretical studies, which are connected with the first practice teaching period - where the comprehensive school pupil, the teacher and the basic modes of interaction between them are in focus - include a course in developmental psychology, a course in observation techniques and a course in general didactics. At best the solution of integration gives student teachers concepts and causes, as well as a more profound vision to continue planning their teaching activities together with supervisors.

5. A concluding remark

I have above briefly described the theoretical background, general principles and some concrete solutions by means of which we have tried to integrate theoretical studies with practice teaching in our primary teacher education programme. The view I have chosen myself omits one important factor, which affects student teachers' reflective thinking and also - as we have noticed - their future teaching behaviour. That factor is the M.A. thesis, which is quite an extensive study unit in the new teacher education programme.

The M.A. thesis also integrates theory and practice in teacher education, but on a more abstract level than the solutions described above. The M.A. thesis in our teacher education programme is a scientific inquiry about a problem arising from the teacher's professional field. The majority of the theses have been empirical inquiries. It has been proposed that the student teacher should choose a theme for her/his thesis from the problems she/he has faced in practice teaching. The idea has worked successfully and connections between theory and practice have been created. Moreover, we have got feedback from school administrators, who consider classroom teachers from the new teacher education programme to be more capable and more willing to start projects in order to develop the school and the professional field of teaching than classroom teachers from the former preparatory programme.

References

Giroux, H.A. & McLaren, P. 1986. Teacher education and the politics of engagement: the case for democratic schooling. *Harvard Educational Review 56*, 3, 213-238.

Hytönen, J. 1982. *Opettajankoulutuksen teoria-aineksia* (Theory Elements in Teacher Education). Helsingin yliopiston opettajankoulutuslaitoksen tutkimuksia n:o 1.

Hytönen, J. 1989. Kasvatustieteen teorian merkitys ohjaustapahtumassa (The use of educational theory in supervising practice teaching). In Ojanen, S. (ed.) *Akateeminen opettaja*. Helsingin yliopiston Lahden tutkimus- ja koulutuskeskus. Täydennyskoulutusjulkaisuja No. 4. 114-122.

Liston, D.P. & Zeichner, K.M. 1991. *Teacher Education and the Social Conditions of Schooling*. New York & London: Routledge.

McLaren, P. 1989. *Life in Schools*. New York & London: Longman.

Peters, R.S. 1977. *Education and the Education of Teachers*. London: Routledge & Kegan Paul.

Zeichner, K.M. 1991. Contradictions and tensions in the professionalization of teaching and the democratization of schools. *Teachers College Record 92*, 3, 363-379.

Educational Knowledge and Reality

Juhani Jussila

The relation between educational knowledge and action is a question which puzzles teachers, students and researchers time after time. Opinions claiming that educational knowledge is too far from everyday affairs are predominant. An important reason for this kind of thinking is that there are several types of knowledge in the field of education. The complexity of practical educational situations requires that we are able to use various types of knowledge appropriately.

In our attempt to understand the various problems related to the educational phenomena and to the nature of educational knowledge, it is useful to start with Karl Popper's classification of reality: we have to deal with three different worlds: world 1 consists of physical objects and organisms; world 2 includes subjective experiences and human consciousness; and world 3 is composed of human language, culture and cultural products. All these worlds must be taken into consideration if we want thorough insight into the process of education, its conditions and its consequences.

In the field of education we need knowledge that refers to singular facts, statistical information and general explanations of things, events and situations. In addition to this, the use of modal knowledge, information concerning possibilities we have in front of us, is also needed. Sometimes, in order to achieve our educational objectives, conditional and instrumental knowledge is needed. At times it is necessary to have evaluative knowledge that tells the significance of an act or thing in relation to some criteria. (Niiniluoto 1989, 54-56.)

The problems of theoretical and practical knowledge occupied philosophers already in ancient Greece. Plato payed attention to 'maker's knowledge' which a handicraftsman needs in making articles. Aristotle layed the foundations of the model of practical inference, which has been in a crucial position in the development of research on human action.

If we want to evaluate the usefulness of educational knowledge, we ought to find out under what kind of scientific conditions and in which situations the knowledge is obtained. In addition to this, we should know how the validity of the knowledge is indicated and what the limits of the validity are. Moreover, we should take into consideration the purpose the knowledge will be used for, who uses the knowledge and what the situation is where one is acting. Educational situations are not regular in the sense that some specific factors could always be causal determinants of human development. Many educational processes depend on how individuals pay attention to various rules and the kind of logic they follow in their actions.

In speaking of educational phenomena, we still must analyze the role of theoretical knowledge and general laws, and, on the other hand, the role of such practical knowledge of how the various individual and subjective dimension of meanings and experiences manifest themselves.

As long as we discuss the usefulness of educational knowledge, it is necessary to take into consideration that separate pieces of knowledge are of little significance in practical situations. Educational knowledge becomes meaningful only when it is connected to some theoretical and situational background. An educator needs an understanding of both theoretical and contextual relations: theory gives concepts which are needed to describe situations in ordinary life and it also gives a frame of reference to the description; it also offers alternative explanations which can make these situations comprehensible.

The relationship between educational knowledge and its use is complicated because we have different kinds of conceptions of the nature of educational theory. The most common view is that theory is a general presentation of knowledge. It specifies the relations among concepts and variables and makes it possible to explain, control and predict phenomena. In this sense a theory gathers together separete pieces of information into a consistent whole. In educational processes we also need a theoretical approach which is more closely connected to the actual situation and the personal thinking and experiences of the educator. Such theory is often called 'user's theory' or 'implicit theory'. It is not as exactly articulated as a scientific theory but this is a matter of necessity owing to the fact that education in practice is always a personal matter, directly or indirectly. It is human interaction.

In speaking of the application of educational knowledge or educational theory we should remember two different meanings of the word 'application':

1) If the question is merely one of a relation which is mechanical, a cause-consequence connection, the basis of action can be a general theory which is applied to some actual situation. Such condition we often meet in technology or natural sciences.

2) Teachers and educators adapt theories in a different way. They use theoretical knowledge of education in their action so that it has an effect through their own

personal reflection. The application of educational theory is not similar with procedures in physical experimentation or in industry: it is not meaningful to try to guide educational processes mechanically as in the paper industry where chemical theories are applied. In education the matter under consideration is not like surgery either, where physiological and medical knowledge is applied in a heart operation. Education is, moreover, not decision-making like in exercise of the law when the court pronounces sentence on a person in a concrete criminal case so that legal security is maximized.

Educational situations are social processes, and their character depends on the engagement, aims and experiential background of all participants. That is why educational theories should be interpreted individually according to each situation.

It is a well-known fact that educational research frequently exploits concepts, theories and methods from other branches of science, e.g. psychology, sociology and philosophy. Interdisciplinary research as such is an important way to produce new points of view, but it easily leads to a pattern of thought where the main point is forgotten. In the study of personality, socialization, or learning we are dealing with matters which are psychological or sociological in character. Therefore, it is quite natural that we have difficulties in the application of such data in educational relations. Education takes a shape of its own and this must be taken into account in the construction of theories. Without the connection of aims, methods and situational factors the knowledge of personality, socialization or learning consists of scattered facts viewing matter from an educational standpoint.

Educational processes are sensitive social events in that their character depends on the engagement, experiences and personal aims of participants. Therefore, educational theories and educational knowledge should be interpreted in a proper way depending on the situtation and factual relationships. That is why it is inadequate to try to deduce technical norms from theories. Instead, we need alternate theories and interpretations of concrete events which make it possible to act rationally in new situations and we need manifold conceptual descriptions of educational reality. Personal experiences together with theoretical reflection make it possible to arrive at a solution which suits the respective educational goal and situation. Educational theories are an auxiliary means to the teacher or educator in that they make it possible to interact with the reality. Theories are also useful in that they help us to find grounds for educational decisions. This means not that an educator should all the time be thinking his grounds and motives, but that he should, when needed, be able to justify his action. (Peterssen 1983, 12.)

In discussing the application of educational knowledge, it is useful to remember that teachers and educators as well as pupils and educands are all both subjects and objects at the same time in educational situations. They all have their personal backgrounds and aims. Theories of education have their influences through the personalities and reflection of the participants. Not all of the problems and possibilities of each situation make their appearence in the theory. They should become evident to the participants themselves in the situation in question. It is a social process which can be understood only in its context, not from the outside.

* * *

The task of educational research is to investigate that goal-oriented process, or interaction, which seeks to advance human growth and which is connected with a certain cultural context. Persons who are engaged in educational research as well as social scientists in general have been confused in recent years as a consequence of Norwood Hanson's idea that we cannot make direct observations of reality: our experiences always depend on what we know of the research object beforehand. Our perceptions of the world of education are therefore dependent on our conceptual frame and previous experience. The uniqueness of educational situations leads also to problems in producing educational knowledge: a method which is suitable in one situation may not work well in other conditions or when employed by other people. The educational process is similar to the continuous change of reality pointed out by Heraclitus 2500 years ago when he said that one cannot step twice into the same river because new water is always running past. The researcher is bound to put the question of how to develop a theory based on events of educational reality which are unique in character, bound to time and place, and where the observations already contain theory.

The aim of science has traditionally been considered the description of reality. Lately, there have been efforts to question this view. Postmodern thinking, which can be seen as an extension of the Frankfurt School, maintains that scientific theories cannot refer to reality; they can only simulate it (Lenzen 1987, 52). Underlying this thought is de Saussure's semiotics, which stresses the arbitrary nature of the sign in language; concepts are not connected with reality, but only to other concepts, and it becomes the aim of science to clarify how people talk about reality. From this foundation the interest in studying education from the viewpoint of linguistic usage is understandable. The research process concentrates on what people say about education and in educational situations as well as on what meanings they attribute to educational matters. The focus of the analysis is discourse, the way of talking. Now we must ask whether it is possible for us to understand education as a whole in this way.

If one thinks that the essence of education is in spoken language and communication, and if the methods of research are borrowed from those of literary critisism, it is natural to consider that the interpretations of discourses manifest and construct reality. Such interpretations can also increase our understanding of how people talk about language. On the other hand, our view of educational reality in its entirety, of how education becomes evident as action in the real world, will remain deficient. Research work which has been founded upon discourses is a kind of "science of conversation". It does not encompass the constraints of the educational process or its practices and consequences, which differ in nature from linguistic usage. Erik Allardt (1992, 2-4) has referred to a similar case in stating that the research on the structure of meanings and the ways of talking is insufficient in the theory construction of social action: it must take into account material conditions, social structures and social institutions as facts.

We certainly need symbols and language in making observations of the various manifestations of the world of education and in establishing educational theories. It is well known, however, that Ludwig Wittgenstein's theory of language as a picture of reality has turned out to be too simplistic. On the other hand, in his later philosophy the notion of a language game as the creator of the connection between reality and language is more comprehensive. Jaakko Hintikka (1976, 110-112) and Aulis Aarnio (1990, 117-125) have presented in their analyses how an expression representing a certain state of reality is determined by the way language is used in that specific situation; this is a clear reference to language games. Although our concepts do not describe reality in a straightforward manner, this does not mean that the link between concepts and reality is lacking. This relationship has a complex character which is revealed by a language game proceeding in accordance with certain rules. That is why we must observe the realm of education more closely, concentrating on what the theoretical foundation of our observations is and what kind of conceptual system and grammar we use in analysing our observations.

The history of educational research has shown that the controversies between various research orientations as to the most suitable approach in educational research are relatively fruitless. Indeed, a qualitative orientation has increased recently in empirical reasearch, and hermeneutical methods now emphasize methodological rigor and reliability (e.g. König 1990, 929). In endeavouring to comprehend to educational reality as a whole, it is necessary to point out the importance of the concurrent use of various theoretical viewpoints and methodical approaches: studying external conditions necessitates a different approach than the elucidation of individual experience and meanings. This is to say that the research has to be done on various theoretical levels, keeping in mind under what conditions and in what respects the information obtained is valid. Therefore, it is even more important than before to be aware of the character of the research object and the theoretical basis of the methods one is using.

Although the process of education is unique and the ways of experiencing it are individual, one must not be led to imagine that the traditional search for truth in educational research is somehow unrealistic. It is necessary to start out from the conception that where education is concerned, the truth has several forms. The knowledge of individual educational situations is unique, as in historical reasearch. There are, however, also regularities in the world of education, and it is possible to explain them by combining information that has been received from different viewpoints. It is thus possible to get an idea of the process of education, with its background conditions and consequences, and also, on the other hand, of the ways of experiencing education and the meanings associated with it. Without any connection to the basic factors of educational reality, without paying attention to its non-linguistic aspects, and without knowing what rules people take into consideration in their actions, writing on educational matters has no foundation.

* * *

The researcher faces a new frontier in trying to describe, understand and explain things which have been previously unknown: This is the border between the researcher and reality. The curtain separating the two - be it composed of observations or language - is not an impenetrable one, however, as Jaakko Hintikka (1988, 164) has aptly stated. This observation recalls the words of Immanuel Kant, who pointed out that observations without concepts are blind and that concepts without observation are empty. In speaking of language, Hintikka refers to relativism, which leads so easily to fruitless incommensurability and internal contradictions.

The discussions of paradigms prompted by Thomas Kuhn have broken down many borders in educational research as well as in many other human and social sciences, although they have also created new ones. I refer here to the kind of borders which have recently been noticeable in methodological discussions. These borders are to be found between various branches of science as well as between different paradigms, between the self and the external world, subject and object, individual and general, systematic observation in standardized experimental situations and participant observation in natural situations.

Particularly problematic in speaking of crossing borderlines are situations where the researcher and the object of study are thought to be so intertwined that one cannot always make out who, the researcher or his subject, is constructing the reality in question each time; this is true especially if scientific research is considered equal to myths, stories and fairy-tales. The situation is the same as that of the psychoanalyst who could not decide if he was therapist or patient. Conceptions referring to the difference between scientific research and everyday beliefs mix with each other in a way that does not credit serious scientific work. It is necessary for this reason to determine when there is sense in crossing the borders in research and when it is a question of getting dressed in the emperor's new clothes as part of a fad.

Education in Finland is currently under the magnifying glass of society's decision-makers, and the position of educational research at the university is being passionately debated. We have good reason, following Knut Tranöy (1976, 287-288), to stress that it is necessary to prove that the resources for scientific research and education are justified, and also that scholars have to be trusted by virtue of the ideals of objectivity, professional honesty and the search for truth.

Reaching this goal requires that one be able to direct attention to the appearance of the social world as both objective and subjective reality simultaneously and that one can distinguish between these two. Peter Berger and Thomas Luckmann (1966) have emphasized this in forming the grounds of social constructionism, which is easily forgotten or overlooked. Also Jean Piaget (1965, 274), who is known as a representantive of individual constructionism, emphasized that he started most decidedly from the viewpoint that the world exists independently of our consciousness and that we analyse it into parts in the interaction of the organism and the environment. Educational research is very deficient if one confines oneself to the meaning people attach to various things and if one does not pay attention to intersubjective reality, the historical and biological conditions of human action, and the ties between the external world and individual consciousness. The border between the researcher and reality becomes problematic in the sense that it is not clear what we mean by educational reality. Numerous different ways for defining the concept of education can be found, and there is also variation in the conceptions of educational reality. The extreme alternative is that one considers it impossible to determine the research object in advance if one does not specify at one and the same time the research method as well. The idea then is that it is the research method in particular which constitutes the object. It is, however, possible to note that it would be a very haphazard journey if one were to select one's route without knowing where one was going. The question is nearly the same as if one could invent a remedy before one has discovered the illness to be cured. As a matter of fact, this problem came up when operational definitions were discussed. It is possible to say that the method used will construe the object, but only in part: the object exists but it gets its structure only when it has been conceptualized. (Morin & Brunet 1992, 107.)

One consequence of beginning to outline the practices of education through discourse was that in the first place attention was paid to the use of language in educational processes. The same research methods which were used by researchers in literature were considered appropriate in education. A strong concentration on linguistic problems was surprising because social interaction is linguistic only in part: love is not primarily talking about love, and fear is also something else than talking about fear. There are many other things influencing educational processes than those which are of a linguistic nature - regardless of the meanings various participants in the educational process give to them - and it is the task of the researcher to ascertain these effects.

In an effort to understand the viewpoints of those who are taking part in the educational process, it is essential that the researcher be part of the evironment which he is studying. This is important because it enables him to give an account of the meanings people give to various matters. It is the same question to which the advocate of the old Frankfurt School, Theodor Adorno (1962), alluded in stressing that the aim of empirical social research is to find out the truth and the untruth of what the phenomena being studied themselves want to be. The understanding of everyday conditions does not imply that the researcher should keep within the bounds of the concepts of the everyday use of language: it is quite possible to use and create theoretical concepts for research, provided that their connection to the social reality is sustained (Kannisto 1986, 175-176).

In speaking of ethnography and of research which is closely related to life, Hammersley (1992, 50-55) has accentuated that the researcher has to maintain his independence: as he makes a statement, it must not change the essential viewpoint of the reality which is the object of study such that it would affect the validity of the statement in that situation. The issue is simply that the results of research are assumed to be founded on conclusions drawn from documented materials in a way which does not depend on the wishes of the researcher. This is something entirely different from the fact that what we know about some phenomenon influences how we conceptualize it.

The proximity of the investigator to his object of study comes out in a pronounced manner in educational action research, where a teacher does research on his own work. There are no obstacles to combining the work of a teacher and that of a researcher, but if a teacher's own wishes and sympathies begin to direct what he considers the results of the study, he has followed the wrong lead. Even though the thought of keeping the roles of teacher and researcher separate, in the sense of independence described above, may seem overwhelming, it is absolutely indispensable if one wishes one's research to seem credible (Hammersley 1993).

The closer the researcher comes to reality, the better chances he has of presenting true statements concerning it. He does not imitate reality, but only construes it conceptually. It is not worthwhile to talk about context-free 'final truths' in connection with research on educational action and society, but neither do we need to sink to the depths of relativism. In trying to understand a phenomenon on the basis of the phenomenon itself, the content of some statements is more relevant and valid than that of others in regard to the phenomenon in question. Reality may never be investigated exhaustively because the educational process proceeds in time and is thus in a continuous state of change; however, this does not prevent us from striving for more sophisticated theories than before: using various criteria, it is possible to investigate the validity of descriptions yielded by different conceptual systems (cf. Niiniluoto 1990, 50).

The association of science and ideology has already been dealt with for a long time in the sociology of knowledge, and particularly in Marxist research and critical philosophy. Educational research and the philosophy of education are without doubt connected with each other such that research needs as a background philosophical ideas of man and society as well as conceptions of nature and the fundamental conditions of education. Just as W.C. Runciman (1969) stated, in his well-known analysis of society and political theory, that the philosophy of politics and the sociology of politics are inseparable, the philosophy of education, with its ideological basis, and the theory of education are closely bound together. We are obliged for that reason to see them as interacting elements. This does not mean that we should consider scientific research and ideology to be identical. They belong to different areas, and they both have their own tasks. It is the obligation of the science of science to analyse how the increase of information concerning society and education affects education in practice and the character of its research in historical perspective.

Habermas' idea of an 'interest in emancipation' has often been used as an argument for uniting research and political action. Rudolf Lassahn (1982, 129) has aptly pointed out that 'interest' in Habermas is a concept on the system level. It refers to a fundamental tendency of mankind and the preservation of the human species: "Technical use, practical understanding and emancipation determine those particular viewpoints from which only we can grasp the reality in itself" (Habermas 1975, 134). Here it is essential that all these three should be taken into consideration together and simultaneously: information widens technical power; interpretation makes it possible to direct the action within the framework of common traditions; and the analysis emancipates the consciousness from a dependency on supposed forces (Habermas 1975, 136). It is therefore astonishing that some people think they are able to proceed from a single knowledge interest in scientific research, e.g. to promote the interests of a certain group of people instead of searching for the truth, and not take into consideration the original unity of knowledge interests which Habermas meant.

Michel Foucault (1966, 314-) was positively right in stressing how modern man is a product of modern thinking and knowledge. As is well known, the same thought is in the background of all planning and developmental work based on scientific research: one thinks that by increasing knowledge we can also create better possibilities for human life. It is possible to see, agreeing with Hammersley (1992, 155 n. 24), a clear difference between scientific research and political action which rests upon the fact that the immediate aim of research is the production of knowledge, while political action aims to direct development. This is not to undervalue the efforts of the sociology of knowledge to scrutinize the link between knowledge, society and power as well as the historical character of knowledge. The issue is what Ilkka Pirttilä (1994) has recently established, i.e. that knowledge and power as social phenomena and objects of sociological research are not one and the same thing: "That a proposition is socially conditional does not mean itself anything in regard to the truthfulness of the knowledge in principle." When the investigator puts forth a statement concerning the object of study, trying to say something informative about the process of education, he comes to make a proposition whose truthfulness can and, indeed, must be analysed.

Those who have become enthusiastic about Foucault's thoughts would do well to take note of the remark of Martin Kusch (1993, 161-162) that Foucault in many cases emphasized that he was not interested in how nature determines our knowledge. According to Kusch, Foucault is doing research on usage and scientific discourses as if nature did not restrict how they are constructed. Kusch is pointing up what is often overlooked, i.e. that Foucault's choice was methodological: he was studying science 'as social community and an extension of clusters of theories' from the viewpoint of the history of science. Foucault could very well treat his subject merely as a discourse, because the objects of his study were scientific texts and the discussion on them. In the study of the educational process, the object is different: in an attempt to ascertain the process which is actually occuring in nature and in a certain society, it is necessary to consider obective facts if one wants to understand education in its entirety. A discursive approach can naturally reveal how some people talk or write about education, but one has to see this only as a part of that process. In educational research, one meets distinctively the conditional basis of the acquisition of knowlege - if for no other reason than the fact that values are incorporated in the aims of education. This is why scholars in the field should take an interest in how people have carried on education in various societies and at different times, how educational processes have been studied, and what has been written about them, for in this way they will grasp the regularities which underlie and structure the process of education.

References

Aarnio, Aulis. 1990. Oikeuskielen tulkinta ja ymmärtäminen. Näkökulma Wittgensteinin myöhäisfilosofiaan. In Hirvonen, Ari & Tuori, Kaarlo (Eds.) *Oikeus, kieli ja kritiikki*. Helsingin yliopiston julkisoikeuden laitoksen julkaisuja D. 9. Hakapaino: Helsinki. 99-127.

Adorno, Theodor W. 1962. (1957) Soziologie und empirisiche Forschung. In Horkheimer, M. & Adorno, Th. W. *Sociologica II. Reden und Vorträge.* Frankfurt am Main: Europäische Verlaganstalt. 205-222.

Allardt, Erik. 1992. Sosiologian kuvauksen ja teorian haasteet Euroopan mullistusten aikana. *Sosiologia 29*. 1, 1-10.

Berger, Peter L. & Luckmann, Thomas. 1966. *The Social Construction of Reality.* Harmondsworth: Penguin.

Foucault, Michel. 1966. *Les mots et les choses*. Paris: Gallimard.

Habermas, Jurgen. 1975. (1968) Tieto ja intressi. In Tuomela, R. & Patoluoto, I. *Yhteiskuntatieteiden filosofiset perusteet. I.* Helsinki: Gaudeamus. 118-141.

Hammersley, Martyn. 1992. *What's wrong with ethnography? Methodological explorations.* London and New York: Routledge.

Hammersley, Martyn. 1993. On teacher as researcher. *Educational Action Research 1*, 425-445.

Hintikka, Jaakko. 1976. Language-Games. In Hintikka, Jaakko (Ed.) *Essays on Wittgenstein in Honour of G. H. von Wright*. Acta Philosophica Fennica 28, 1-3. Suomalaisen Kirjallisuuden Kirjapaino: Helsinki. 105-125.

Hintikka, Jaakko. 1988. Onko totuus saavuttamaton. In Kusch, M. & Hintikka, J. *Kieli ja maailma.* Oulu: Pohjoinen. 143-169.

Kannisto, Heikki. 1986. Ymmärtäminen, kritiikki ja hermeneutiikka. In Niiniluoto, I. & Saarinen, E. *Vuosisatamme filosofia*. Porvoo etc.: WSOY. 145-243.

Kusch, Martin. 1993. (1991) *Tiedon kentät ja kerrostumat*. Oulu: Pohjoinen.

König, Eckard. 1990. Bilanz der Theorieentwicklung in der Erziehungswissenschaft. *Zeitschrift für Pädagogik 36*, 6, 919-936.

Lassahn, Rudolf. 1982. *Einführung in die Pädagogik.* 4. Aufl. Heidelberg: Quelle & Meyer.

Lenzen, D. 1987. Mythos, Metapher und Simulation: Zu den Aussichten systematischer Pädagogik in der Postmoderne. *Zeitschrift für Pädagogik 33*, 1, 41-60.

Morin, Lucien & Brunet, Louis. 1992. *Philosophie de l'education. 1. Les sciences de l'education.* Sainte-Foy (Quebeck): Les Presses de l'Université Laval.

Niiniluoto, Ilkka. 1989. *Informaatio, tieto ja yhteiskunta. Filosofinen käsiteanalyysi.* Helsinki: Valtion painatuskeskus.

Niiniluoto, Ilkka. 1990. *Maailma, minä ja kulttuuri.* Helsinki: Otava.

Petersen, W.H. 1983. *Lehrbuch Allgemeine Didaktik.* München: Ehrenwirt.

Piaget, Jean. 1965. *Sagesse et illusions de la philosophie.* Paris: P.U.F.

Pirttilä, Ilkka. 1994. Tiedonsosiologinen maailmanmalli. *Tiedepolitiikka 19*, 1, 43-48.

Runciman, W.G. 1969. *Social Science and Political Theory.* 2nd ed. Cambridge: Cambridge University Press.

Tranöy, Knut Erik. 1976. (1972) Tieteenharjoituksen ideologia. In Tuomela, R. & Patoluoto, I. *Yhteiskuntatieteiden filosofiset perusteet. II.* Helsinki: Gaudeamus. 268-294.

The *Deutsche Didaktik* and the American Research on Teaching[1]

Pertti Kansanen

0. Introduction

My interest in the *Didaktik* began in my early studies in psychology of education and learning theories. All the textbooks were in English or in Swedish and the students of my generation got a very thorough understanding of the American way of thinking of the educational problems and how to do research correctly. There was only one way: according to the method of science. In the late 1970s I found the book of Wolfgang Klafki *Studien zur Bildungstheorie und Didaktik* in a book sale. I remember that I understood practically nothing of its content.

As a university teacher of foundations of education one of my courses was about the basics of *Didaktik*. It was always confusing to use the concept of *Didaktik* without really knowing what it meant. I knew that it came from Germany but its content was from the American curriculum research or from American educational psychology. Although we co-operated with the IPN in Kiel when the curriculum research was at its peak in the 1970s the content of the *Didaktik* were in the background. At that time there was no need to get acquainted with the human sciences or *Geisteswissenschaften* or its method, hermeneutics.

Gradually, when the general attention began to focus more and more on the theoretical background of the empirical models, the question of the nature of the *Didaktik* became of current interest. In the Finnish teacher education *Didaktik* is the main subject and because my chair represents teacher education, it became a personal problem to find an answer to the question *What Didaktik really is*?

It was not possible to get an answer from the American literature or from the German literature of curriculum research. After some conceptual analyses there was no other way to solve the problem than to begin to read German *Didaktik* books, among others the old Klafki. But it was not easy at all. In the Nordic university libraries you cannot find a sufficient number of German books, you must go to Germany. Luckily, in those German universities that I know the libraries are excellent. This literature opens a wholly new world and you can notice how it is possible to think differently of the same problems.

---

[1] Paper presented at the NERA Annual Meeting in Vasa, March 1994

1. The background of German didactic models

The German *Didaktik* (*didactica*) was founded by Wolfgang Ratke and Johan Amos Comenius (1592-1670) at the beginning of the 17th century. Its idea was to develop a general method to teach compared with the logical method which at that time was thought as the best way to present the teaching content in order to bring about learning. *Didaktik* was a practical and normative doctrine by nature (*Lehrkunst*) and the best-known presentation of its early characterisation is *Didactica Magna* by Comenius.

The position of *Didaktik* with regard to pedagogics (*Pädagogik*) changed during the next centuries. The work of Johann Friedrich Herbart (1776-1841) brought again the status of *Didaktik* to the centre of education with his formal stages and with his principle of education through instruction. At his time *Didaktik* had a strong position as a science of education (*Wissenschaft*). *Didaktik* was mainly concerned with education at schools. Schools were practically the only places were organised education took place.

At the beginning of the 20th century *die Reformpädagogik* acquired its great representatives (Kerschensteiner, Gaudig, Petersen) with the main focus on child-centered activities. On the theoretical side, the pedagogical thinking was dominated by *geisteswissenschaftliche Didaktik* (Nohl, Weniger, Klafki) until the early sixties when the empirical-analytic paradigm gained some ground (Heimann, Schulz, Otto). Thirdly, critical-communicative *Didaktik* offered an alternative based on critical theory and especially on the ideas of Jürgen Habermas.

In addition to these three theoretical models, in the contemporary *Deutsche Didaktik* there are numerous minor variations and local versions. The development has brought the main models closer to each other as the theoretical background of the models has been analysed. *Die Didaktik* has also been in close contact with teacher education. By nature, *die Didaktik* in Germany has always been philosophical thinking, theorising, and construction of theoretical models.

2. The American tradition of research on teaching

The American tradition of research on teaching and on the problems of curriculum development is not as long as in Germany. It can be traced back to pragmatism and to its main representatives Herbert Spencer (1820-1903) and William James (1842-1910). The influence of John Dewey and William Heard Kilpatrick in particular has been great. At the same time, educational psychology, with Edward L. Thorndike, achieved its central position in research on teaching in the US. The fundamental interest in teaching was practical by nature.

The predominant approach to the problems of teaching has been research on teacher and teaching effectiveness. Along these lines there has been a series of model building from Mitzel, Dunkin & Biddle to Shulman (Gage 1963, Dunkin & Biddle 1974, Shulman 1986a). The purpose of this kind of thinking has been an attempt to find those teachers who could attain the best possible results and to determine those factors which are crucial in planning and acting in the teaching process. In this way research is also connected with teacher education.

On the theoretical level, the development of theoretical models has concentrated on empirical research and on testing these in real situations. Mastery learning in particular, based on the ideas of John B. Carroll and Benjamin S. Bloom, is well known in this respect. Bruce Joyce and Marsha Weil have collected the various philosophical and psychological strategies and formulated applications to teaching.

Most of the research on teaching has been empirical and with quantitative methods. The latest developments, however, have led to alternative approaches. The most commonly used process-product -paradigm is not as dominant as earlier and studies with qualitative research methods have greatly increased. The theoretical background of the discussion of research problems is clearly research methodology which leads to other philosophical questions.

3. Some preliminary conclusions

3.1. Comparison of terminology

Die Didaktik - didactics

Didactics (*die Didaktik*) is a difficult term to use. Its origin is an applied translation from Greek meaning both teaching and learning (Heursen 1986, Knecht-von Martial 1985). Wolfgang Ratke and Johan Amos Comenius were the founders of this terminology and the first to launch the term in their writings, but *Didaktik* was an artificial term in a certain way. The respective family of words in Greek was not translated with a German word but with an application of the original *didáskein* via Latin *didactica*, which also was artificial. Its very first meaning was about the same as the art of teaching or *Lehrkunst*.

Die Didaktik was gradually taken into more general use alongside *die Pädagogik* or pedagogics, but its use was limited to German-speaking countries or to countries having cultural relations with Germany. As a result, *Didaktik* is nowadays in use in Central Europe and in the Scandinavian countries, but it is practically unknown at least in English or French-speaking countries in the area of education. The very word can be found in dictionaries with quite different meanings, however. It is based on *didascalia* in the meaning of didactic poem (Blankertz 1975, 14), and that makes its use most awkward and disturbing.

What is then the proper translation of *Didaktik*? (cf. also Kansanen 1987). If we emphasise the normative side of *Didaktik*, the most convenient word would be the art of teaching. This expression, however, already has its own context (cf. Gage 1978). But if we want to keep the definition as wide as *Didaktik* is nowadays, the art of teaching is too narrow because there is no reference to learning in its meaning. Naturally, there are various interpretations of teaching (cf. Smith 1987 with older references), but what is essential in this analysis is that *Didaktik* is at the same time a second order term. It is thus a model or a system of how to envisage the teaching-learning process as well as a kind of metatheory where the various models can be compared with each other.

If we emphasise the descriptive side of *Didaktik*, the research aspects come to the centre of its content. Then the proper word would be research on teaching. If we look into the well-known research models of Mitzel, Dunkin & Biddle, and Shulman, the aspects of learning can also be seen there. The difference between the descriptive *Didaktik* and research on teaching is in their background or in the purpose of their model building. The first, *Didaktik*, is mainly meant for teacher education and the models are based on a philosophical conception of man and on the nature of research concerning his education. The empirical research results are not a prerequisite for its building, but the results are used, naturally, in a corrective way when they are in conflict with the model variables. The second, research on teaching, is meant for research purposes and that's why the models are mainly inductive by nature and based directly on research results. The practical conclusions can be drawn from these models and thus they can function in teacher education, too.

Earlier, the American research mainly based on empiric-analytic foundations, most of the research was conducted with the so-called process-product model. The picture has changed and alternative research paradigms can be found (cf. Guba 1990). At the same time, the philosophical foundations have become more versatile and the situation reminds us in many ways of the respective state of affairs in Germany.

Looking at the same problem from the other side, we can pose the question of how to translate *research on teaching*. Here we can find an easy solution: it is *Unterrichtsforschung*. This translation makes it clear that the core is classroom research with a psychological or social psychological emphasis. The research problems are mainly empirical. If we now compare *Didaktik* and *Unterrichtsforschung* we can notice essential and great differences in their use. First of all, *Unterrichtsforschung* is only part of *Didaktik* and with their comparison the different philosophical traditions come to the fore. Secondly, *Didaktik* is of genuine German origin. It is based on philosophical tradition of its own with such names as Kant, Herbart, Schleiermacher etc. The different schools of *Didaktik* which exist in the German literature mainly refer to the German tradition. It should be noted that the more empirical elements in a model of *Didaktik*, the more references can be found to American research on teaching. The content of *Unterrichtsforschung* consists of empirical results; it is descriptive by nature and it is classroom research employing all possible means and in principle with different kinds of philosophical backgrounds.

I have also suggested that *Didaktik* can be found in the textbooks of educational psychology (Kansanen 1987). Those books (e.g. Gage & Berliner 1984) have lengthy sections containing background material of a purely psychological nature, as well as clear normative sections. The psychology of education and *Didaktik* are linked together, being referred to as educational psychology. Teaching methods in particular are those parts in which the practical side comes into consideration. The theoretical references are to the theories of curriculum and that is why the analysis of the term *Didaktik* is not possible without considering the meaning of curriculum.

Curriculum - das Curriculum - der Lehrplan - die Didaktik

Josef Dolch (1959, 318-319) has pointed out the early use of the word *curriculum* in both German and English. In Anglo-Saxon educational literature it has remained since then in the terminology, in German it was displaced during the 18th century with the word plan and further with teaching plan (*Lehrplan*). It was the philantropists who took the new term up and Herbart was already using it at the beginning of the 19th century.

The word curriculum came back into use in German during this century, in the late sixties (e.g. Blankertz 1975, 118-122). Through American influence, *das Curriculum* was taken into use as a better version of a teaching plan. It was Saul B. Robinsohn (1967) who introduced a new approach of curriculum planning with his book which at the same time was broader in its meaning than the former teaching plan (*Lehrplan*). The application of the term curriculum was based on the American idea of *Reformpädagogik* by John Dewey and its focus was on every individual pupil and his learning experiences. Herwig Blankertz describes (1975, 122) the differences between these two terms from the German point of view. The teaching plan had become more and more a plan for the teacher of how to organise the activities when teaching a special subject and choosing the content within this subject. The new conception of the teaching plan *curriculum* concentrated on every pupil and his learning.

Thus, the curriculum was defined through the learning experiences, and common to various definitions was the focus on the individual pupil and the learning experiences which he was to encounter during his time at school (cf. Hosford 1973). If we take the broadest meaning of the curriculum, it consists of all the experiences organised during the time the school is responsible for the pupil. This also contains, by definition, such experiences which are not consciously planned but which are happening in the school. Thus, in this case there is no room left for the hidden curriculum because all the experiences are within the curriculum. (cf. Jackson 1992, 4-12.)

Gradually, the meaning of curriculum was broadening and as curriculum theory, its scope was nearly the same as traditional *Didaktik*. The word, *das Curriculum*, was directly taken into use without any special translation and its content was becoming more and more the same as *Didaktik* with a particular emphasis of its own (cf. Frey 1971). Wolfgang Klafki (1974) wrote an article in a dictionary under the common heading "Curriculum - Didaktik" and it seemed that *Didaktik* would be subsumed under the more general curriculum. It was a radical interpretation of traditional *Didaktik* and it showed a certain change in thinking about the old subdiscipline of education. It was, however, only a question of how to compare these two aspects which were parts of the more general *Didaktik*. In this article Klafki described the old directions of didactic models and in addition to that, the aspects of curriculum planning and controlling or evaluation. So one can say that it reflected at least a different conception of the problems of *Didaktik* and it had great influence on practical curriculum development.

The research on curriculum problems concentrating on development, planning, and evaluation grew greatly during the 1970s and it reached its peak in the early 1980s. The results were reported in large handbooks (Frey 1975, Hameyer & Frey & Haft 1983). *Didaktik* and curriculum theory were considered as parallel areas of the same subdiscipline. During these years the emphasis was on curriculum theory and it had a very important role in the efforts to achieve school reform, and in particular in reforming the old teaching plans into a modern curriculum.

It is not easy to define the curriculum, and difficulties arise because curriculum as a concept has numerous semantic contents and nuances depending on the context in which it is found and on the purpose for which it is used. Reisse (1975) points out that the term curriculum is strongly culture-bound which is why comparison of its meanings across linguistic boundaries is fraught with a variety of difficulties. Additionally, of course, any term may also have several meanings within a specific cultural environment (cf. Connelly & Lantz 1985). The American influence of the implementation of the term curriculum can be evaluated from the point of view of planning and evaluation of education in institutes. The problems of formulating educational goals and objectives as guidelines for teaching practice were focused on, and methods of evaluation, both in the classroom and on the school level, became more important than earlier.

The question of the relation between *Didaktik* and curriculum has gradually lost its interest and the status quo seems to have been achieved. The impulses have come from the American research, but there is hardly any evidence of impulses in the opposite direction. One could conclude that the didactic aspects of curriculum have integrated into *Didaktik*. Zimmermann (1986) is of the opinion that discussion can be reinstated because we now know the good and bad sides of the problem.

3.2. The independence of education as a discipline

The first independent chair of education was established at the University of Halle in 1779. The very first professor of education was Ernst Christian Trapp (1745-1818). His idea "*Versuch einer Pädagogik*" was to no longer base education on philosophy and theology but on the nature of man and on contemporary society. He also spoke about such modern research methods as observation and experience as a basis for conclusions. This professorship is considered as the start of an independent discipline and it is clear that it happened in Germany where there had been much educational thinking in the area of philosophy and theology. It took about one hundred years before independent professorships in education were established in England, Scotland and the US (cf. Sjöstrand 1967, Wulf 1977.)

From the beginning, education was considered as an independent discipline with its own problems. The current classification of education can be traced back to the German tradition and there are certain differences between the German and the American way of classification. There are three or four common basic problem areas: education in general, the psychology of education and sociology of education. Usually, the classification must be made according to one criterium at a time, and this point can arouse some confusion. The most common criterium is the classification of disciplines. However, there can be such criteria as the content of education or the age of pupils etc. In a well-known German example (Röhrs 1969) general education consists of pedagogics and *Didaktik* (*Pädagogik und Didaktik*) and the latter is usually seen as a subdiscipline concentrating on the questions of teaching. General education is further divided into sub-areas using educational reality and the period of life as criteria for the division. This leads to school education (*Schulpädagogik*), special education (*Sonderpädagogik*), pre-school education (*Pädagogik der Frühen Kindheit*), vocational education (*Berufspädagogik*) and adult education (*Erwachsenenbildung*). In addition to the basic classification, the history of education and comparative education overlap all the other areas.

In British educational literature there has been a consensus of opinion about the nature of education. However, Paul Hirst does not agree with the term *discipline*, he prefers to use the term *a field of study* (Hirst 1983; Tibble 1966). Lee S. Shulman also says the same: "... education is not itself a discipline. Indeed, *education is a field of study*" (Shulman 1988, 5). So there are some doubts about the status of education depending on the way we think of formulating its problems. At the same time, there are many aspects and many possible approaches resulting in various research methods which have their foundations in several background disciplines. That is why any attempt to make a systematic classification does not succeed without many simultaneous criteria.

In any case, in British as well as in American educational literature, the sub-area of *Didaktik* seems to be lacking. As we have seen earlier, much of its content belongs to educational psychology. In the American literature of research on teaching, the problems of teaching and learning in general are usually held together without any theoretical model building. Attention is paid to the methodological problems, and there the various background principles can be seen. In German educational literature, didactic problems define an independent subdiscipline of education which really is quite the same as general education, however, with its own point of view. The area of *Didaktik* is mainly larger than educational psychology and it includes much philosophical and theoretical thinking. In German literature *Didaktik* and educational psychology are clearly separate fields with different representatives. The situation in Great Britain and the US is quite the contrary; the same people are working in this common area. Naturally, there are differences as to the importance given to some aspects of the problems, e.g. the role of learning in the teaching process.

4. How the traditions separated

We know that at the end of the 19th century American educational research had many contacts with German research. Walter Doyle (1993) refers to the term *didactics* as he quotes Paul Woodring's text in the 1975 Yearbook of the National Society for the Study of Education with the astonishing remark that a chair of didactics was created at the State University of Iowa in 1873. We also know that John Dewey was a member of the first executive council of the *National Herbart Society* that later changed its name to *the National Herbart Society for the Study of Teaching* and once more to *the National Society for the Study of Education*. The texts of Hegel and Herbart were known to him and to other colleagues through translations. At the personal level there were numerous contacts and study trips and consequently the language of education was common to both.

Stephan Hopmann has analysed in depth the early history of German *Didaktik* and the common background of German and American *Didaktik* (Hopmann 1992). Although there were many contacts with Herbart's ideas and progressive education had its respective version as *die Reformpädagogik* in Europe, the contacts suddenly ended at the beginning of the 20th century. Hopmann (1992, 7) also remarks that there were, however, certain differences between the progressive movement and reform pedagogy; the latter emphasised teacher education and schooling whereas the former was more concerned with social change and politics.

Further, Hopmann (1992, 8) states that it was the educational psychology that the Americans (Hall and Dewey) took from Herbart, not the whole of *Didaktik*. The main reasons for the spread of *Didaktik* in Germany were the state guidelines for the curriculum and the system of teacher education in the seminaries (Hopmann 1992, 4-5; Hopmann & Riquarts 1992, 22). These required a central solution and central models of schooling problems in society. The criticism of Herbartianism that was a mechanical application of the ideas of Herbart led to the reform pedagogy and through it to new conceptualisations of *Didaktik*. This new development did not reach American education and at the same time the word didactics disappeared from the terminology. That means a different sort of development in both countries with amazingly great disparities.

Walter Doyle and Ian Westbury (1992, 138-145) explain the development of American education by referring to the structure of governance in the system of schooling. The local boards of education had the responsibility for the effectiveness of the schools and the role of the superintendent was central. The interaction between the school and the local community was very intimate. Although the individual states had constitutional responsibility, the control was merely a formality. In addition to these basic characteristics, the absence of church had many consequences in the curriculum and in practice. The model of teaching was the same as in business life: "They (teachers) were and are a labor force to be motivated and managed as any large enterprise's labor force was motivated and managed." (Doyle & Westbury 1992, 140). It is easy to see, I think, that the atmosphere was not very encouraging to independent and autonomous action. Accountability was always narrow and the local boards and public held a direct control over the school and the teachers.

Instead of *Didaktik*, psychology of education took its place as a discipline of the science of education in the US. At the same time this line of research in Germany became separated from *Didaktik*, although there was at first a close relationship between them. Concening this development in the US there is a certain important point that needs special attention. Doyle and Westbury (1992, 141) quote Ellen Lagemann as saying "one cannot understand the history of education in the United States during the 20th century unless one realizes that Edward L. Thorndike won and John Dewey lost". This can be seen e.g. in the well-known textbook of Robert M. Travers (1978) where Thorndike's position is central. Afterwards it is easy to say that this way of thinking was too fragmented and its behavioural and experimental features were too narrow to apply to the whole process of education. This phase, I think, however, was necessary in the development of educational research. The defects are not to be found in the psychology of education itself but in the way it was applied over the whole field of education without alternatives.

Empirical research can be done in many ways. Some of us do it without thinking of the philosophical assumptions behind the procedures. Some practical problem guides the thinking, and research methods are selected according to their practical value in finding solutions to the problem. In this example the awareness of the method has not aroused and the way of doing research is self-evident and it is not problematised. To follow the Kuhnian language, the action is happening inside the dominating paradigm where all researchers agree with each other. I think that looking at the problems of education through the glasses of psychology of education has been this kind of paradigmatic work and all the participants have been content with it. The science of education has been a practical tool in administration at the local level and attention has been on practical problems in real situations. Thinking with psychological concepts is thinking with the problems of students (learning, motivation, ability, achievement, tests etc.). It is at the same time empiric-analytical as well as democratic towards the process of education.

The other side of the coin, many American colleagues claim, is that the practical approach has neglected the importance of content in the curriculum and instruction. Naturally, psychology of education as a background discipline leads thoughts to the psychological content and particularly to management and learning problems. These are no doubt an important part of the totality but not sufficient in themselves. Finding the content has led to looking at the European *Didaktik* again but this time from a special point of view, *Fachdidaktik*. I dare to point out, however, that the psychological problems have not vanished from the instructional process and that's why the general aspect of *Didaktik* should be kept in mind constantly.

Peter Menck (1993) has referred to the early German tradition of empirical educational research that was existing alongside the old tradition of *Didaktik*. As we know, Wilhelm August Lay and particularly Ernst Meumann (1862-1915) are its main representatives. Meumann had been a student of Wilhelm Wundt but his interest had turned to the problems of *Didaktik*. Their experimental *Didaktik* could not gain status and it got only a marginal position in the area of German *Didaktik*. Heinz-Elmar Tenorth (1988, 214-219) calls it "der szientifische Flügel" - the scientific wing of reform pedagogy. Although its influence seemed to be small it had some very important disciples who were to continue the approach in a way that was discovered only after many years. Aloys Fischer (1882-1937) was the first and he turned the research from experimental to descriptive and Peter Petersen (1884-1952), a disciple of Meumann, was the other one who is generally considered the founder of the so-called *Pädagogische Tatsachenforschung*, empirical research on pedagogical facts.

Fischer developed his ideas in a phenomenological sense but independent from Husserl (Tenorth 1988, 217). The basic idea in this descriptive empirical research was to look at the instructional process as a phenomenon that is as much as possible theory-free. That requires observing the process as it is, without any predetermined theoretical assumptions. Petersen developed a sort of observation system in his Jena-Plan-School. The most important and central concepts were the pedagogical situation and the various aspects, pedagogical facts, that describe the pedagogical situation.

The descriptive line of *Didaktik* did not succeed in gaining a respected academic position and it remained a side trend behind the erudition-centered *Didaktik*. The latest well-known work is that of Friedrich Winnefeld in Halle (1957).[2]

This line of development of the descriptive *Didaktik* is the German alternative to the empirical *Didaktik*. The literature is almost unanimous in stating that educational psychology in Germany has been an independent discipline without any close relations with *Didaktik* and that the empirical influences have in general come from the US and from its psychology of education. In *Berliner Didaktik* the empirical approach is to be seen but gradually that part diminished with the work of Schulz. The contacts between German *Didaktik* and American research on teaching have been rather few.

As a conclusion it can be said that the erudition-centered *Didaktik* did not gain a footing in the USA in the beginning of this century. Instead, the reflection on teaching continued in psychology of education. In Germany reform pedagogy transformed into erudition-centered *Didaktik* which got later some rival directions. The empiric-analytical approach did not succeed in getting a break-through in Germany in spite of a good beginning with Meumann and Lay. It lived some time as descriptive *Didaktik* but it did not develop into psychology of education. The latter got its impulses from the USA and has been a separate area alongside *Didaktik*.

The work of Peter Petersen is, however, very interesting from the viewpoint of German-American relations. Herman Röhrs (1993, 11-19) takes Petersen as an example from this interaction as he analyses progressive education in the USA and its influence on European reform pedagogy. On the practical level the discussion about progressive education was international and the well-known systems of school reforms of Helen Parkhurst, Carleton Washburn, Maria Montessori, and Peter Petersen were influenced by each other. "New Education Fellowship" was a connecting link between educational practical workers and researchers. In 1928 Petersen made a visit to the USA and became familiar with the American situation. This was later seen in his Jena-Plan. In spite of these kinds of relations between individual colleagues, the main trends in the area of *Didaktik* grew apart. Naturally, the political situation in the world contributed, but there were some ideological reasons, too.

---

[2] Of special interest is that Toivo Laurilehto wrote a dissertation in 1980 "Sykliketjut opetuksessa" (Verbal compound cycles in the classroom teaching situation) following the method of Winnefeld

5. Some contemporary trends

The role and the meaning of knowledge in educational research in recent years have clearly increased. Some interesting viewpoints have been brought into the terminology and communication. Shulman (1986b) has focused on teachers' understanding of the subject they are teaching. In addition to content knowledge, the essential substance is pedagogical content knowledge. This same aspect has been referred to by different names, and Reynolds (1992, 5) introduces various alternatives which all have something special: content-specific pedagogy, subject-specific pedagogical knowledge, content-specific cognitional knowledge, and subject matter specific pedagogical knowledge. (cf. also Gudmundsdottir & Shulman 1987, 54-55.)

This old idea of pedagogical reduction of factual content for the purposes of teaching is known in the traditional German *Didaktik* as *Fachdidaktik*. The modern view of *Fachdidaktik* takes into consideration all the factors in the teaching-learning process from the content point of view. It was Ch. Helwig who as early as 1619 made a distinction between the common aspects of teaching (*didactica generalis*) and the content aspects of teaching (*didactica specialis*) (Knecht-von Martial 1985, 17-28). The very idea, however, in spite of the use of different language, has always been known to parents and teachers. In any case, this comparison between pedagogical content knowledge and *Fachdidaktik* could offer useful knowledge to both sides.

The renaissance of content has aroused the idea of comparing the erudition-centered *Didaktik* with the new conception of research on teaching. Stephan Hopmann (1992) as well as Gudmundsdottir and Grankvist (1992) have already made a start in this respect. The latter also tell that the new trend had nothing to do with the European *Fachdidaktik* (1992, 185). Although the idea looks the same there are essential differences.

Looking at Shulman's content knowledge and pedagogical content knowledge it is clear that the focus is on the substance that is the content in the instructional process. The introduction of these concepts has brought about lively discussion in the journals. When looking at this discussion more precisely, it is possible to notice that the focus is not exactly on the substance or subject-matter but on the structural analysis of this substance. What is presented is a reflection on what kind of elements there may be in the specific content. Frank Achtenhagen (1992, 316) remarks e.g. that "...the distinctions are useful: knowledge is regarded as the "subjective" aspect of subject matter and content as "objective"." My suggestion, however, is that the presentation of content is as formal and general as the former focus on students' properties: learning, motivation, achievement, etc. In other words, psychology of education still has a strict hold but from a different point of view than before and the possible paradigm shift is only a change in the themes and topics.

The change is, however, to be seen in the area of curriculum planning and in emphasising the importance of the instructional content in the curriculum. This is to do with cognitivism and action research along with the growing power of the teachers themselves in preparing their own curriculum. But if we compare the pedagogical content knowledge with *Fachdidaktik* on this level we soon notice that there are different kinds of assumptions behind them.

There are, however, only slight principal differences between pedagogical content knowledge and *Fachdidaktik* or between content knowledge and *Fachwissen* if we compare the German models of *Didaktik* with the American way of thinking in this respect and leave the erudition-centered *Didaktik* out of this comparison.

Heimann, Otto, and Schulz had, in principle, in their *Berliner Didaktik* a very similar conception of the position of content in the curricular or in the instructional process as Shulman. Because their starting point was empiric-analytical there was no exact standpoint according to the substance but only a category named. Content was one central category in the totality of their model and the criteria of selection were brought from developmental psychology and the life situation of the pupils but no direct stand was taken on the selection of subject-matter. Later with the changes produced by Schulz, the model got much of the same characteristics as erudition-centered *Didaktik*. With these changes the position of content changed as well.

Critical-communicative *Didaktik*, however, has a clearly normative overstructure where content is selected with certain value criteria. The same features are found in critical pedagogy in the US but content in this model is not reflected from the viewpoint of structural analysis. The background is openly normative and political, and this is to be seen also in the instructional process itself. Group work and co-operation are the slogans, but the nature of pedagogical content knowledge is general and does not focus on the school subjects as much as on the methods.

The curriculum movement brought its own conception of content with educational aims, goals, and objectives. It was structure again that was the guiding principle. Taxonomies stimulated very precise analyses of the psychological content. They also offered a good basis for the presentation of subject matter, but this movement had weaknesses in other respects and that's why it was not possible to build a curricular totality with this idea. The same can be said of cybernetic *Didaktik* although the level of exactness required was extremely high. Content was given in the curriculum and the method algorithm was based on the conditions of the factors given in the curriculum.

In all these examples the common aspect is the interpretation of content as formal and general that can be further refined in the curriculum and in the teacher's work. This is very understandable because the models are built for all possible situations, subject matter and curricula. The selection of content is left to practitioners, textbook writers, and curriculum makers. The researchers have stayed out of this process because the concept of doing research has not included taking a stand on value questions and schooling policy. Changing the theme to the content of the teacher's thinking or to the cognitive structure of the teacher's thinking does not change the basic assumptions of the research; it remains within the same paradigm. Naturally we get other types of research results and our attention is focused on other kinds of problems, but the philosophy of doing research stays on the same foundation.

Gradually I am coming to my point of how to compare the German *Fachdidaktik* with the American way of thinking about the same problems. The comparison can not be made by putting content or the analysis of the structure of this content or the typical characteristics of this content side by side. This is only on the surface. The various curricula or textbooks can be compared in this way, of course, but if the motive is to compare the whole frame of reference, this is not enough. In the German *Didaktik* the key is German idealism with such names as Immanuel Kant, Friedrich Schleiermacher and Wilhelm Dilthey with many more recent names. We know this as *Geisteswissenschaften* and I have used erudition-centred *Didaktik* as its English equivalent in the area of teaching. The whole comparison can be simplified by putting erudition-centered *Didaktik* on one side and all other models on the other.

One more point must be taken into consideration and it is the purpose for building the *Didaktik* models and the models for research on teaching as well as the conception of curriculum planning. It has already been said that the German *Didaktik* models are built for teacher education and for instructional planning. The various models have a different approach to the selection of content and especially to the normative criteria of this selection. Only the erudition-centered *Didaktik* has a clear stand on this question and that is why it has a close connection with *Fachdidaktik* or pedagogical content knowledge. Research models are general and take no position on value questions, and that it why content comes into research according to its position in the design.

In the erudition-centered *Didaktik* the main task is seen as a theory of educational content (Theorie der Bildungsinhalte). The content of education is selected according to its value in the curriculum and in the instructional process. The decision is always based on tradition and history. It is also dependent on the particular group of students for whom the curriculum is written. As can be seen, the erudition-centered *Didaktik* has its main role in the planning and writing of the curriculum where the decision-making is openly value laden. The selection of content that is at the same time the selection of aims and goals is, however, not pure policy making because the erudition-centered *Didaktik* claims to have educational autonomy and expert knowledge in educational matters. In this system there are both formal educational criteria and clear normative decisions.

Another side of this question is that the same decision-making continues inside the curriculum when teachers select the instructional content or the textbook writers decide on what is valuable to be transformed from content knowledge to pedagogical content knowledge. This second part is similar to teachers' work in general and in this phase the problem of learning comes to the fore. Erudition-centered *Didaktik* has been criticised for its neglect of learning and method problems in the instructional process. It has been more interested in what is valuable in content and what is worth teaching than controlling how much has been learned. In this respect there has been development in recent years.

6. The Nordic Alternative

The *Didaktik* in the Nordic countries has been educational psychology with an emphasis on the teacher and on the instructional process. The German *geisteswissenschaftliche Didaktik* has been practically unknown with certain, mainly Danish and Norwegian exceptions (e.g. Reidar Myhre, Torstein Harbo and Bjørg Gundem). When the educational psychology line and the *geisteswissenschaftliche* line get into contact with each other there are almost always conflicts to be seen. Yet the focus of both approaches is the instructional process, teaching and the teacher, and the curriculum etc. Why is it so?

You can easily notice this conflict if some researcher is asked to evaluate the works of the other trend. The representative of the empirical research quite often says that it is not research at all, it is a number of opinions. The hermeneutic says that empirical research is only making notes about something which already exists in practice, but what then. Quite often they speak of technology, that means thinking without creativity or alternatives.

A very good example of this situation is Wolfgang Brezinka who is said to represent critical rationalism along with Karl Popper's ideas. He divides education into three parts: philosophy of education, education and the practice of education. The first, philosophy of education is not scientific at all. It is policy making, decision making, opinions etc. Naturally you must have some basis for your opinions but that does not change the essence of it. The practice of education is action and has nothing to do with science or *Wissenschaft*. You can use facts behind your practice but the action itself is not scientific by nature. Only the description, understanding, and explanation of the educational process is scientific.

Consequently, the difference is not in the focus, in the instructional process itself, it is behind the process in its theoretical assumptions. And it is not possible to combine them, the conflict remains.

In the Nordic countries with the above mentioned exceptions the instructional process has been investigated along the empirical paradigm That is why it is very difficult to make a difference between *Didaktik* and educational psychology. In practice these two subdisciplines have been a combined area with certain emphases on partly one, partly on the other. If someone has claimed the name of *Didaktik* in his writings, it has not been the *geisteswissenschaftliche Didaktik*. A good example of this has been the *Didaktik* discussion in Sweden.

In general, we can note two perspectives in this discussion. The first line of research concentrates on the macro level, on the societal, economic and political prerequisites of education (Dahllöf, Lundgren, Englund). We can not say that it does not take the very process into consideration, because Dahllöf and Lundgren has made this kind of research, too. Its emphasis and interpretation of the empirical results has, however, been on the macro level, on the frames. Curriculum research is a natural part of this line.

The other line of research concentrates on the other end of the educational process, on the learning of individual students or on their conceptions of this learning (Marton, Svensson, Lybeck, Kroksmark). There is much research in this group of the very instructional process but, nevertheless, the focus is on a certain part of. So one could say that both of them have a very important part of the instructional process as its focus but neither of them can be said to concentrate on the whole totality of the instructional process. And that is the very essence of the general *Didaktik*. At the same time this object is enormously large and that is why most of the research is done in some subarea. Accordingly, the totality of *Didaktik* is divided into subdisciplines and naturally research made in these areas is *didaktikal* research, too.

It is not exceptional to have various schools of thinking inside the *Didaktik*. As a matter of fact, it is more a rule to have different approaches of *Didaktik*. There is, however, a big difference on what bases they are considered different. In the German *Didaktik* it is the decision of the philosophical background which is determining the different perspective. In the Nordic countries, I think, the differences are not seen through these kinds of lenses although the philosophical base may be different. The various approaches live inside the empirical tradition although there may be a strong emphasis on e.g. phenomenology. Naturally this leads to the comparison of *Geisteswissenschaften* and phenomenology which is not an easy task. In any case the starting point has been within the empirical tradition and the various emphases have emerged gradually alongside the research work.

7. A Concluding Remark

The erudition-centered *Didaktik* is a very good example of how education is a national thing in a broader international context. The question, however, remains whether it is possible to compare educational systems in different cultures and to transform new ideas from another culture if there is not criticism enough.

REFERENCES

Achtenhagen, Frank. "The Relevance of Content for Teaching-Learning Processes." *Oser - Dick - Patry 1992*, 315-328.

Blankertz, Herwig. 1975. *Theorien und Modelle der Didaktik* (9., verarbeitete und erweiterte Auflage). München: Juventa. (1.ed. 1969)

Connelly, F.M. - O. Lantz. 1985. "Definition of Curriculum." *Husén - Postlethwaite 1985*, 1160-1163.

Dolch, Josef. 1959. *Lehrplan des Abendlandes. Zweieinhalb Jahrtausende seiner Geschichte*. Ratingen: Aloys Henn.

Doyle, Walter. 1993. "Curriculum Theory in the Preparation of Teachers." *Paper Presented at the Symposium: Didaktik and or Curriculum*, Kiel, IPN, 5-8 October 1993.

Doyle, Walter - Ian Westbury. 1992. "Die Rückbesinnung auf den Unterrichtsinhalt in der Curriculum- und Bildungsforschung in den USA." *Bildung und Erziehung 45*, 137-157.

Dunkin, M. 1988. "Comparative Studies in Teaching and Teacher Education." *Postlethwaite 1988*, 60-66.

Dunkin, M. (Ed.) 1987. *The International Encyclopedia of Teacher and Teacher Education*. Oxford: Pergamon Press.

Dunkin, Michael J. - Bruce J. Biddle. 1974. *The Study of Teaching*. New York: Holt, Rinehart and Winston.

Frey, Karl. 1971. *Theorien des Curriculums*. Weinheim: Beltz.

Frey, Karl (Hrsg.) 1975. *Curriculum - Handbuch I-III*. München: Piper.

Gage, N.L. 1963a. "Paradigms for Research on Teaching." *Gage 1963b*, 94-141.

Gage, N.L. 1978. *The Scientific Basis of the Art of Teaching*. New York: Teachers College Press.

Gage, N.L. - David Berliner. 1984. *Educational Psychology*. Third Edition. Boston: Houghton Mifflin.

Gage, N.L. (Ed.) 1963b. *Handbook of Research on Teaching*. Chicago: Rand McNally.

Guba, Egon G. (Ed.) 1990. *The Paradigm Dialog*. Newbury Park: Sage.

Gudmundsdottir, Sigrun - Lee S. Shulman. 1987. "Pedagogical Content Knowledge: Teachers' Ways of Knowing." *Stromnes - Sovik 1987*, 51-83.

Gudmundsdottir, Sigrun - Rolf Grankvist. 1992. "Deutsche Didaktik aus der Sicht neuerer empirischer Unterrichts- und Curriculumforschung in den USA." *Bildung und Erziehung 45*, 175-187.

Gumpert, Edgar B. (Ed.) 1990. *Fit to teach: Teacher Education in International Perspective*. Center for Cross-cultural Education. College of Education, Georgia State University. Atlanta, Georgia.

Gundem, Bjorg B. 1992. "Didaktik in Skandinavien." *Bildung und Erziehung 45*, 189-200

Gundem, Bjorg B. 1992. "The Place of Didactics in Curriculum in Scandinavia." *Paper Presented at the Annual Meeting of the American Educational Research Association, AERA, April 20-24, 1992* San Francisco.

Gundem, Bjorg B. 1992. "Notes on the Development of Nordic Didactics." *Journal of Curriculum Studies 24*, 61-70.

Haller, Hans-Dieter - Hilbert Meyer (Hrsg.) 1986. *Ziele und Inhalte des Unterrichts. Enzyklopädie Erziehungswissenschaft, Band 3*. Klett-Cotta.

Hameyer, Uwe - Karl Frey - Henning Haft (Hrsg.) 1983. *Handbuch der Curriculumforschung*. Weinheim: Beltz.

Harbo, Torstein. 1986. "Didaktiska teman i vårt århundrade." *Harbo - Kroksmark 1986*, 49-95.

Harbo, Torstein - Tomas Kroksmark. 1986. *Grundskolans didaktik*. Lund: Studentlitteratur.

Heursen, Gerd. 1986. "Didaktik, allgemeine." "Fachdidaktik." *Haller - Meyer 1986*, 407-415, 427-439.

Hirst, Paul (Ed.) 1983. *Educational Theory and Its Foundation Disciplines*. London: Routledge & Kegan Paul.

Hopmann, Stephan. 1992. "Starting a Dialogue. Roots and Issues of the Beginning Conversation between European Didaktik and the American Curriculum Tradition." *Paper presented at the Annual Meeting of the American Educational Research Association in San Francisco, April 1992.*

Hopmann, Stephan - Rudolf Künzli. 1992. "Didaktik-Renaissance." *Bildung und Erziehung 45*, 117-135.

Hopmann, Stephan - Kurt Riquarts. 1992. "Didaktik - didaktikk - didactics." *NORDEUROPA-forum Nr.1, 1992*, 21-24.

Hosford, Philip L. 1973. *An Instructional Theory: A Beginning*. Englewood Cliffs: Prentice-Hall.

Husén, Torsten - T. Neville Postlethwaite (Eds.) 1985. *The International Encyclopedia of Education*. Oxford: Pergamon Press.

Jackson, Philip W. 1992. "Conceptions of Curriculum and Curriculum Specialists." *Jackson (Ed.) 1992*, 3-40.

Jackson, Philip W. (Ed.) 1992. *Handbook of Research on Curriculum*. New York: Macmillan.

Jaeger, Richard M. (Ed.) *Complementary Methods for Research in Education*. Washington, DC: AERA.

Kansanen, Pertti. 1987. "The Curriculum as a Factor Directing Actual Teaching." *Malinen - Kansanen 1987*, 47-68.

Knecht-von Martial, Ingbert. 1985. *Geschichte der Didaktik. Zur Geschichte des Begriffs und der didaktischen Paradigmen*. Frankfurt am Main: R. G. Fischer.

Klafki, Wolfgang. 1974. "Curriculum - Didaktik." *Wulf 1974*, 117-128.

Lazerson, Marvin. 1990. "Research and Teacher Education in the American University." *Gumpert 1991*, 65-84.

Malinen, Paavo - Pertti Kansanen (Eds.) 1987. *Research Frames of the Finnish Curriculum*. Research Report 53. Department of Teacher Education, University of Helsinki.

BEST COPY AVAILABLE

Menck, Peter. 1993. "Some Remarks on Research in Education and Didactics in Germany." *Paper Presented at the Symposium: Didaktik and or Curriculum*, Kiel, IPN, 5-8 October 1993.

Oser, Fritz K. - Andreas Dick - Jean-Luc Patry. 1992. *Effective and Responsible Teaching.* San Francisco: Jossey-Bass.

Patry, Jean-Luc. 1992. "Didaktik und Curriculum. Konfrontation - Koexistenz - Komplementarität." *Bildung und Erziehung 45*, 213-223.

Postlethwaite, T. Neville (Ed.) 1988. *The Encyclopedia of Comparative Education and National Systems of Education.* Oxford: Pergamon Press.

Reisse, Wilfried. 1975. "Verschiedene Begriffsbestimmungen von "Curriculum": Überblick und Ansätze zur Präzisierung." *Frey 1975*, 46-59.

Reynolds, A. 1992. "What Is Competent Beginning Teaching? A Review of the Literature" *Review of Educational Research 62*, 1-35.

Robinsohn, Saul B. 1967. *Bildungsreform als Revision des Curriculum.* Neuwied: Luchterhand.

Röhrs, Hermann. 1969. *Allgemeine Erziehungswissenschaft.* Weinheim: Beltz.

Röhrs, Hermann. 1993. "Die Reformpädagogik (Progressive Education) in den USA und ihr Einfluss auf die reformpädagogische Entwicklung in Deutschland." *Pädagogische Rundschau 47*, 3-27.

Shulman, Lee S. 1986a. "Paradigms and Research Programs in the Study of Teaching: A Contemporary Perspective." *Wittrock 1986*, 3-36.

Shulman, Lee S. 1986b. "Those Who Understand: Knowledge Growth in Teaching." *Educational Researcher 15 (2)*, 4-14.

Shulman, Lee S. 1988. "Disciplines of Inquiry in Education: An Overview." *Jaeger 1988*, 3-58.

Sjöstrand, Wilhelm. 1967. *Education As An Academic Discipline.* Acta Universitatis Upsaliensis C. 12.

Smith, B.O. 1987. "Teaching." *Dunkin 1987*, 11-15.

Stromnes, Åsmund L. - Nils Sovik. 1987. *Teachers' Thinking. Perspectives and Research.* Tapir.

Tenorth, Heinz-Elmar. 1988. *Geschichte der Erziehung. Einführung in die Grundzüge ihrer neuzeitlichen Entwicklung.* Weinheim: Juventa.

Tibble, J.W. 1966a. "The Development of the Study of Education." *Tibble 1966b*, 1-28.

Tibble, J.W. (Ed.) 1966b. *The Study of Education.* London: Routledge & Kegan Paul.

Travers, Robert M.W. 1978. *An Introduction to Educational Research.* Fourth Edition. New York: Macmillan.

Twellmann, Walter (Hrsg.) *Handbuch Schule und Unterricht*. Band 8.1. Düsseldorf: Schwann.

Wittrock, Merlin C. (Ed) 1986. *Handbook of Research on Teaching*. Third Edition. New York: Macmillan.

Wulf, Christoph. 1977. *Theorien und Konzepte der Erziehungswissenschaft*. München: Juventa.

Wulf, Christoph (Hrsg.) 1974. *Wörterbuch der Erziehung*. München: Piper.

Zimmermann, Werner. 1986. "Ist die Stunde des Curriculum vorbei? Wege und Irrwege der Curriculumentwicklung in der Bundesrepublik Deutschland." *Twellmann 1986*, 107-121.

Concepts of Freedom in Danish School Legislation

Sven Erik Nordenbo

1. In the history of Danish primary and secondary school (*"Folkeskolen"*) the concept of 'freedom' is normally associated with two remarkable dates. The first one is the year 1814 when universal compulsory schooling was introduced. The 1814 Act - which in fact consists of several Acts, cf. *Royal decrees of July 29th 1814* - requires that all children reaching the age of seven should attend school up to the year of confirmation, i.e. seven years of school attendance, but at the same time the school is 'free' precisely in the meaning that it is free of charge, a 'non-fee-paying school' (cf. Larsen, 1893, 253-79; Nellemann, 1966, 82-88; Markussen, 1988, 258-298. This principle of freedom, in the sense of 'a school free of charge' has afterwards been generally recognised by the Danish state for the entire educational system. Consequently, all education, including universities and other higher education, are today free of charge in Denmark.

2. The second date is not so familiar to the ordinary Dane although the matter is common knowledge. It is the year in which the Danish Parliament passed a law that later became known as the law on 'school-freedom' or on 'free-schools'. The year is 1855, only six years after Denmark introduced constitutional monarchy with a freely elected parliament.

This law can, in the language of the time, be described as a transformation of 'universal schooling with compulsory attendance' into 'a legal obligation to educate every child'. The fact was that the law of 1814 on compulsory schooling had been attacked from two sides: for the great majority of the population, the peasant population - but also now and then from the manufacturers, cf. Markussen (1989) - the law appeared as an intervention into their right to dispose of the child's labour. Consequently, they kept the children away from school to a certain extent, particularly in periods in which the farmers were in need of manpower. The authorities' reaction was 'mulcting', i.e. fining parents and farmers who did not observe the law about school attendance. 'Free schooling' was, therefore, considered as 'school coercion'.

A certain resistance also originated from the middle classes in the cities. They wanted fee-paying schools for children of well-to-do-parents and charity schools for the children of parents who did not pay taxes. The fee-paying schools were to contribute to keeping the social classes apart and to giving extended instruction in Mathematics, Geography, History and Modern Languages. Through the 1830s and 40s there was a debate on whether the state or the parents are responsible for the child's schooling. In the first Danish democratic Constitution of 1849 the spokesmen for state and private schooling entered into a provisional compromise by giving sect. 90 in the Constitution the following wording: 'Children of parents unable to finance their schooling, will receive schooling free of charge in peasant schools'. Only the needy should expect to be provided for in the new democratic state. The Constitution of 1849, thus, stands by the same definition of 'free school' as 'school free of charge' as the 1814 Act.

The expected appeasement of the controversy on public versus private schooling did not take place but was fed even more from another quarter: religious parents wanted to give their children a religious education different for the official one. Similar religious movements from the same period are known in other countries resulting in private schools of local interest. In Denmark the development is, however, different due to the fact that the parents here are borne along on a wave of both religious and national revival which encompasses more than school matters. It is in this context N.F.S. Grundtvig (1783-1872) and Kresten Kold (1816-1870) enter into the picture. What brings them together is a common wish to revive the whole Danish people religiously and nationally by means of preaching and instruction which in these years find new patterns in *narrative*, *informal speeches*, and *songs*. Grundtvig produces the necessary texts which Kold implements in his schools.

The 'Peasant Folk High Schools' and 'Free Schools' which are inaugurated in the 1850s and 60s influenced by Grundtvig-Kold do not focus on traditional school teaching. They form part of a revival that also comprehends the church and the social organisation. The 'Free-School Act' of 1855 exempts the children from compulsory school attendance and thus makes the existence of these schools possible. With the Act of 'school-freedom' or 'free-schools' - the term 'free' is now employed in another sense than in the 1814 Act - it was made legal, so to speak, to 'liberate' schooling from the state. The 'Free-School' Act of 2 May 1855 gave parents and guardians the option of educating children outside the public school as long as these pupils sat for a semiannual examination in the state school. Some few years later, in 1864, the free-school also got the right to conduct the exams themselves (cf. Skovmand, 1978).

The 'Free-School' Act is the law that confirms the legal obligation to educate, but not an obligation for children to attend school, as the central principle in Danish school legislation (cf. Bodenstein, 1982; Engberg, 1986). Although the Act was a concession to the influence of Grundtvig-Koldean views on home education, the Act in no way limits the right of other interest groups to found schools. It is probably this feature that has attracted most attention abroad, i.e. that with the 'free-schools' a tradition has been created for a school structure in which private and public school are not perceived as competitors but as mutually complementing and inspiring organisations which guarantee the right of the individual, both of the child but in particular of the parents, to have a say in the content and methods of the actual school teaching. A particular feature of the Danish system is that, probably unlike the majority of other countries, the founding of a school as either a private school, a 'free-school', or a 'little school' is largely not curtailed by the economic capacity of the initiating parents, as the state, on certain conditions, funds a substantial part of the economy of new schools. In a way it is an extension of the 1814 Act's conception of 'freedom' to include also the free-schools.

The most remarkable feature of the Danish free-schools is, therefore, that in form and content they become expressions of fundamental national movements and needs in the population at a specific time. The character of the individual school is determined more by the commitment which a group of parents brings to a cause, whether religious, educational or political, than by a self-seeking endeavour to secure one's own child a better position in the job-market. In that sense the school becomes a school for life and not for the labour market.

This view prevails in the Danish school system in a degree that it has influenced also many traits of the public school, cf. the following quotation from a recently published book instructing parents how to create a curriculum for their own *public* (state) school:

> A curriculum is a legal description of what a pupil has a right to learn. It is only of interest in marginal cases. It is worth noting that neither parents nor teachers nor pupils find the curriculum as a legal document of any importance. The important thing is whether the teaching is existentially appropriate. Not only: do I learn what I need in order to pas the exams? Or: do I learn what I need to cope with *gymnasium*? But: do I learn what I need as a human being in this society, in this life? (Held, 1991, 19).

It is thus possible to claim that the spirit that about 140 years ago determined Danish school legislation has inspired the whole school system - private as well as public.

3. It is comparatively more straightforward to link the conception of 'freedom in the Danish school' with the 1814 Act's conception of 'school free of charge' and the 1855 Act's conception of 'free-school' than with the following probably less well known as well as more controversial conceptions of 'freedom'.

The 1814 Act determined that the supervision of the Danish school should be the job of the Danish state church. Not until the Danish Social Democrat Party came to power in 1929 did this situation change. In 1933 the Social Democrat Party at length put an end to the church's supervision (*Lov om Folkeskolens Styrelse*, 1933; see also Markussen, 1971). With later changes, led to the provision in the 1975 Education Act for Religious Knowledge to be a subject that gives information on several religions rather than simply preaching the gospel. Thus it no longer makes sense to be exempted from Religious Knowledge, although this option is still in force even in the latest Education Act (*Lov om folkeskolen*, 1993, Sect. 6. subsection 2).

4. Four years later, in 1937, the now 123-years old Act of 1814 was replaced by legislation (*Lov om Folkeskolen*, 1937) which, among other things, introduced a new type of school: the 'examination-free secondary school'. Although later on the Act was characterised as an educational failure (cf. Skovgaard-Petersen, 1978, 49) which in sect. 3 'unfortunately' uses exactly the expression 'examination-free' (cf. Kruchov, 1985, 142), it gives a very true expression to the prevailing aversion to another form of coercion in schools, that is exams. In the Act's sect. 13 this reluctance is expressed in the following way: 'Except for secondary schools' teaching for examination, the Primary and Secondary School (*"Folkeskolen"*) does not end with an examination'. Exams support class society and contempt for manual work, said the Social Democrat Party. The new type of school should rehabilitate manual and practical work and should make it possible for each individual to develop on their own terms, released from previously inflexible examination requirements (cf. *Undervisningsvejledning*, 1942; *Betænkning*, May 1952). Therefore, the school had to be 'examination-free'. The examination-free secondary school was in existence only up to the School Act of 1958 but the idea of a reduction or even abolition of any kind of exams in the Danish Primary and Secondary School has, particularly in the 1970s, ever since left its mark on the Danish school. Today marks exist only in the 8th to 10th forms as an indication of the pupil's performance in the individual subjects, whilst exams have been limited to the Final School Examination at the end of the 9th form. The main impression is therefore a school keeping examinations to a minimum. The Danish Primary and Secondary School are not 'examination-free', but nearly.

5. A departmental order (*Bekendtgorelse om Maalet for Folkeskolens Undervisning*, 1941) granted the teachers in the Primary and Secondary School a formal right to decide themselves which teaching methods to apply in classes, a principle known as 'method freedom'. It was later defined as the freedom for each teacher 'to choose his own methods and teaching aids, and to prepare the lesson in the way he considers most appropriate' (*Undervisningsvejledning*, 1960, 127), provided that the prescribed objectives are attained. This freedom can be interpreted in three ways:

One can first point to the fact that throughout this century a process of professionalisation has taken place within the teaching profession. This is a consequence of improved training but is also the result of an endeavour to strengthen the teachers' role as competent public servants. Therefore a teacher is now expected to be able to prepare his work independently within an indicated framework (cf. Fibæk Laursen, 1976, 67-81 & 84-106). This official expectation has on the other hand been internalised by the teaching profession as a claim to power and freedom in their work. And their profession and work is to provide for the best possible teaching, based on professional competence, with no interference of laymen.

But it is also possible to see 'method freedom' as a right, even for public servants, to express oneself on public matters. Thus 'method freedom' forms part of the Danish teacher's perception of his work: the more or less systematically ordered and justified conceptions he makes about the foundation, means, aims and conditions of his occupation. The development of this principle, i.e. the independent attitude towards and interpretation of the task of the school and the professionally justified selection of content and practices, is determined and is limited in the final analysis by 'what the political public opinion or its representatives consider as reasonable interpretations' (Nielsen, 1980, 20).

Finally, one can consider 'method freedom' as a particular administrative principle which in an educational context has been described as 'licensing' (cf. Goodson, 1988; Haft & Hopmann, 1989). Licensing in curriculum work implies a distinction between curriculum development at a political level and curriculum work in the educational context. Teachers are made responsible for implementing the aims, objectives and content of the school as these are expressed in the curriculum. It is why one speaks of 'licensing' the teachers, implying a moral commitment rather than a legal binding obligation. Teachers sometimes interpret this as meaning that it is up to each teacher freely to choose how they want to carry out their duties within given frameworks ('method freedom'), but in fact the content has already been decided on their behalf (cf. Gundem, 1993, 36). A less friendly but probably not correct way of describing this feature is to claim that Danish teachers conform to the system (cf. Striib, 1989).

The teaching profession's traditional adherence to the principle of 'method freedom' has in the last twenty years been on the decline. The principle fitted easily into the prevailing conception of didactics, in Denmark known as DLH-didactics (Danmarks Lærerhøjskole = The Royal School of Educational Studies, Copenhagen) of which Professor Carl-Åge Larsen was the strongest exponent. This didactics was inspired by the well-known distinction in the German *Geisteswissenschaftlische Didaktik* of the 1950s and 60s between 'didactics in a more restricted sense' and 'the theory of teaching methods' (in German: *Methodik*) and claimed that 'teaching methods' are a matter for the professional teacher. Criticism in the 1970s and onwards of this didactic conception and the growing theoretical acceptance of the view that no factor in the didactic field can be considered as independent of others, means that the view of the teacher as the only one responsible for the applied teaching methods necessarily appears as less convincing. If decisions about teaching methods are no longer a professional matter but have to be related to the political framework and the aim of the Danish Primary and Secondary School, then the teacher cannot any longer make up his own mind about teaching methods.

Finally, a new factor has made an appearance with the Education Act of 1975 and - the latest one - of 1993. Since the 1920s and 30s the influence of progressive education on the day-to-day life of Danish schools (cf. Nørgaard, 1977; Henriksen & Nørgaard, 1983) has made it a truism that pupils should have a say on what goes on in the classrooms. In the Education Act of 1975 this right was codified for the first time in sect. 16. subsect. 4: 'The detailed planning and adjustment of the lesson, including the choice of organisation, methods, and matter, should as far as possible take place in a collaboration between teacher and pupils'.

The tradition of parental influence on the daily life of the school dates back, as mentioned above, to the previous century (cf. Nissen, 1986). From the year 1989 this has been formally established much more strongly by the creation of governing bodies for each school. In these governing bodies (cf. *Lov om folkeskolen*, 1993, sect. 42) the parent representatives have absolute authority over the life of the school. The governing body determines the principles for the running of the school, the organisation of the teaching, the teachers' work schedule, the school budget, approves teaching materials and aids, works out the curriculum, and nominates the school leader and teachers (sect. 44).

With these two codified provisions on pupil participation in the individual lessons and parental management of the individual schools - in defiance of loud protests from teachers, cf. Krogh-Jespersen, 1989 - a situation in which the teacher decides teaching methods independently has been replaced with a situation in which more persons, through a democratic negotiation process, have to enter into a compromise to make the school function.

6. It is debatable whether the creation in the Education Act of 1903 of a unified school system, replacing separated systems, should be considered as given some sort of 'freedom' to attain higher education for the children of the lower classes. In the interpretation of the concept of 'freedom' that has been applied above it will, however, be more appropriate to describe the 1903-Act as expressing a concern for 'social equality' rather than for 'social liberty'.

7. *Concluding summary*: It can be said that 'freedom' has been ascribed to different phenomena through the 180 years' existence of the Danish Primary and Secondary School: It has been interpreted as (1) free of charge, (2) free-schools; (3) 'sermon-free', (4) 'examination-free' and (5) 'method free'. The development of these freedoms or liberties has been accompanied by a movement from centralised rule, through local management, until now the day-to-day running of the school is organised almost entirely by grass-roots management, with more and more individuals having a legal right to intervene, ranging from Parliament, the Minister of Science and Education, the Civil Service, the local authorities, to teachers, parents and pupils. The result has been a school in which the democratic way of life has been an intrinsic goal. Equality has become the major virtue and the will through discussion to enter into a compromise has become a condition for the survival of the school. Many Danes consider these to be the features of a humane school. But another path could have been chosen. Compared with school systems in other countries it appears that the Danish way also has its costs.

REFERENCES:

a.

Anordning for Almueskolevæsenet på Landet i Danmark, af d. 29.juli 1814.

Anordning for Almueskolevæsenet i Kobstæderne i Danmark, Kobenhavn undtagen, af d. 29.juli 1814.

Lov indeholdende nogle forandrede Bestemmelser i Anordningerne om Skolegangen og Udskrivningen af Skolen af 2. maj 1855 ("Friskoleloven")

Lov om Folkeskolens Styrelse og Tilsyn og Tilsynet med private Skoler, Lov nr. 211 af 20. Maj 1933.

Lov om Folkeskolen, Lov nr. 160 af 18. Maj 1937.

Bekendtgørelse om Maalet for Folkeskolens Undervisning, Nr. 242 af 24. Maj 1941.

Undervisningsvejledning for Den eksamensfri Folkeskole, udarbejdet af Det af Undervisningsministeriet under 27. Juni 1941 nedsatte udvalg. Copenhagen: Gjellerup, 1942.

Betænkning fra fællesudvalget vedrørende den eksamensfri mellemskole og landsbyskolen. Copenhagen: Maj 1952.

Undervisningsvejledning for folkeskolen. Betænkning afgivet af det af undervisningsministeriet under 1. september 1958 nedsatte læseplansudvalg. Betænkning nr. 253, 1960 ("Den blå betænkning").

Lov om folkeskolen, Lov nr. 313 af 26.juni 1975.

Lov om folkeskolen, Lov nr. 509 af 1993.06.30.

b.

Eckhard Bodenstein (1982): *Skolefrihed in Dänemark - Studien zur Entstehung eines schulpolitischen Prinzips*. Tønder: Selskabet for Dansk Skolehistorie.

Chresten Kruchov (1985): Socialdemokratiet og folkeskolen. Vejen mod den udelte skole 1900-1984, in: C. Kruchov *et al.* (eds): *Bidrag til Den danske skoles historie 1898-1984*. Copenhagen: Unge Pædagoger, 127-159.

Hanne Engberg (1986): Sådan begyndte det, in: L. Kjær Nielsen *et al.* (eds): *Fri skole - en levende tradition*. Jelling: Dansk friskoleforening, 17-25.

Per Fibæk Laursen (1976): *Læreren fra degn til lønarbejder. Læreruddannelse og didaktik*. Roskilde: Roskilde Universitetsforlag.

I.F. Goodson (1988): *The Making of Curriculum. Collected Essays*. London: Falmer Press.

Bjørg Brandtzæg Gundem: *Mot en ny skolevirkelighed? Læreplanen i et sentraliserings- og desentraliseringsperspektiv*. Oslo: Ad Notam Gyldendal 1993.

Henning Haft & Stephan Hopmann (1989): State-run curriculum development in the Federal Repuplic of Germany. *Journal of Curriculum Studies*, 21 (2), 185-190.

Finn Held (1991): *Skolens egne læseplaner*. Copenhagen: Gyldendal.

Holger Henriksen (ed) (1980): *Skolens opgave - åndsfrihed og demokrati: ni artikler til undervisningslære og pædagogik*. Copenhagen: Munksgaard.

Spæt Henriksen & Ellen Nørgaard (1983): *Vanløse-dagbogen - en reformpædagogisk praksis*. Copenhagen: Gyldendal.

Kirsten Krogh-Jespersen (1989): Skal forældrene lede skolerne? in: Jens Arnum *et al.* (eds): *Hvem bestemmer? En debatbog om ledelse i folkeskolen*. Copenhagen: Gyldendal, 106-118.

Joachim Larsen (1893): *Bidrag til den danske Folkeskoles Historie 1784-1818*. Copenhagen: Thaning & Appel.

Ingrid Markussen (1971): Folkeskolens tilsyn gennem 150 år - en orientering, *Årbog for Dansk Skolehistorie*, 29-43.

Ingrid Markussen (1988): *Visdommens lænker. Studier i evevældens skolereformer fra Reventlow til skolelov*. Odense: Landbohistorisk Selskab.

Ingrid Markussen (1989): Hvad skete der med skolen i 1814?, in: T. Kampmann *et al.* (eds): *Et folk kom i skole 1814-1989*. Copenhagen: Danmarks Lærerhøjskole og Undervisningsministeriet ved Institut for Dansk Skolehistorie, 1989, 35-48.

Aksel H. Nellemann (1966): *Den danske skoles historie*. Copenhagen: Gjellerup, 1966.

Bent Nielsen (1980): Om lærerens metodefrihed og ytringsfriheden, in: Holger Henriksen (ed): *Skolens opgave - åndsfrihed og demokrati. Ni artikler til undervisningslære og pædagogik*. Copenhagen: Munksgaard, 17-34.

Gunhild Nissen (1986): Forældreret - en trossag, in: L. Kjær Nielsen *et al.* (eds): *Fri skole - en levende tradition*. Jelling: Dansk friskoleforening, 85-93.

Ellen Nørgaard (1977): *Lille barn, hvis er du? En skolehistorisk undersøgelse over reformbestræbelser inden for den danske folkeskole i mellemkrigstiden*. Copenhagen: Gyldendal.

Vagn Skovgaard-Petersen (1978): Lighed i skolen - mulighed eller illustion? in: Ingrid Markussen & Ib Goldbach (eds): *Danske skoleproblemer - for og nu*. Copenhagen: Gjellerup, 43-54.

Roar Skovmand (1978): Hvordan de frie grundtvig-koldske skoler vandt indpas i lovgivningen, in: Ingrid Markussen & Ib Goldbach (eds): *Danske skoleproblemer - for og nu*. Copenhagen: Gjellerup: 1978, 31-42.

Andreas Striib (1989): De lovlydige lærere - de lovlydige skoler, in: Jens Arnum *et al.* (eds): *Hvem bestemmer? En debatbog om ledelse i folkeskolen*. Copenhagen: Gyldendal, 80-92.

The Denial of Change in the Process of Change: Systems of Ideas and the Construction of National Evaluations[1]

Thomas S. Popkewitz

Since an OECD report on educational policy in Norway in 1988, The Royal Norwegian Ministry of Church, Education and Research has been developing a model of evaluation. An initial international meeting was held in Oslo to discuss the theoretical and critical aspects of an evaluation of an educational system ongoing major changes in its governance structures (see. e.g., Granheim et al., 1990). I attended that meeting and a subsequent meeting in 1994 that discussed the progress of the evaluation. This essay reflects some of the issues that were raised in that second meeting as I listened to the Norwegian evaluation community present their work. While starting with issues raised in the Norwegian evaluation, however, I use "familiar" examples from the U.S. and, sometimes, from Sweden, to explore the problematic of school evaluation. The choice of examples is pragmatic -- I have a more varied understanding of its research communities and the politics of knowledge that are central themes of my discussion.

As was commented on at my first meeting about The Norwegian evaluation, it is comprehensive in scope, and methodologically diverse (Popkewitz, 1990). Its dialogue continually makes apparent the social democratic traditions of Norway as an horizon for the evaluation. There is a commitment to a participatory process which juxtaposes the normative commitments in a democracy with the development of a scientific expertise for assessing how the goals of schooling are met. There is discussion of the state as a "frame factor" which imposes external restraints on the practices of schooling. At the same time, there are discussions of cooperation between school and parents, and the need for teachers to develop a higher degree of professionalism in curriculum and monitoring of students' achievement. The need to develop technical expertise to collect information at the national level, in fact, is continually seen in relation to various local and professional constituencies.

In this essay, I consider the construction of evaluation systems in Norway, and in other countries, as responses to a crisis in which the particular systems of administration to regulate and discipline liberty no longer work. In one sense, we can think of the current State evaluations as historically emerging from a modernist project to administer schooling. The construction of State evaluations was to order, regulate and give direction to the changes occurring. But the regulating was not only of institutions; the modernization project was to construct dispositions, sensitivities and awareness from which individuals govern their "selfs"; such as how individuals were to think of themselves and act as healthy citizens and as productive members of a society. Evaluation, in this context, can be thought of as more than acquiring information: it was a practice to administer and regulate freedom in a longer term trajectory of what Norbert Elias (1978) called "the civilizing process".

An important strategy within modernity has been a social engineering approach to social change. State policy makers would design policies to solve the social questions of the time and turn to the policy sciences to help craft wiser policies where there was ignorance. Change was to evolve around universal sets of rules which were applied in their uniformity. Evaluation was one practice in directing social change. Today, we can consider that notions of social engineering are no longer viable in the manner that they were once proposed if we consider the problems of regulating that are inscribed in state steering of schools in a decentralized system -- the new sets of problems of regulating are embodied in what sociologists have called variously high modernity, and post-modernity, among others.[2] The shift in the social conditions and shifts in problems of regulating is an unspoken horizon of the Norwegian evaluation.

My intent in this essay is to tie an analysis of the changing conditions to the problem of a national evaluation. Four themes are emerged from the discussion of Norwegian evaluation which I explore through examining practices in the U.S. and Sweden. I do move to these other contexts as examples of issues that, I believe, are cross national boundaries. The themes are:

(1) Changes in the arts, the epistemologies of the social sciences, in economy, and politics have produced new problems of regulation that are also embodied in schooling. I argue that the changes have implications to the ways in which people are to locate themselves in their worlds as active, competent and successful individuals. The changes provide an horizon to the changes inscribed in the evaluation.

(2) Drawing on research in multiple countries, I argue that the systems of ideas in evaluation "deny" the changes through the concepts that are

applied. This "denial" occurs through certain rhetorical strategies which symbolically link past democratic commitments in schooling to current practices, but fail conceptually to make problematic the current relations formed in the school reforms. My examples of this denial relate to concepts of participation and centralization/decentralization which, I argue, are bound to a theory of the State built with the expansion of its welfare functions after World War Two.

(3) My third theme relates to a *doublet* of evaluation: the opening of new spaces for teachers to have greater autonomy through professionalism and participation in schooling also produces new sets of enclosures that exclude. This is a difficult argument as we tend not to think of school practices as simultaneously including/excluding because of the strong normative commitments to schooling in a democracy. But the two co-exist as the distinctions and differentiations of teaching, school subjects, and measurement of achievement constructs boundaries about what is permissible and not-permissible.

(4) The concluding section focuses on the issue of epistemic drift as certain types of problem-solving are sanctioned for school success. In the current moment, this is important when considering the State resources given to testing and measurement in the current evaluation.

An assumption of this essay is what linguistic and social theories have helped us understand: ideas have no essential meaning but are "made sense of" in their relation to other ideas as well as to the historical contexts in which ideas are situated. *Further, the significance of evaluations is not only in the information given but through its systems of classification that provide boundaries about what is to be questioned, organized, and sought as solutions for improving schooling. My strategy in posing the four themes is to inquire into the historical condition of the evaluation through which we reason and ask questions about schooling; thus setting the problem of evaluation as a self-reflectivity that does not take-for-granted either the social conditions in which schooling occurs nor the concepts that select and interpret those conditions.*

EVALUATION AND "THE CIVILIZING PROCESS": THE ARTS, EPISTEMOLOGIES, ECONOMICS, AND POLITICS

Our democratic commitments lead us to consider as important the Norwegian evaluation's attention to the local management of schools, to teachers being received as professionals; and, on the other hand, for the need of the State to monitor the overall performance of its school system to ensure that a just and equitable system is achieved.[1] The dual qualities of a democratic system are assumed as the effectiveness of the evaluations systems is considered. Questions are asked, at one level, about whether school-based evaluations help teachers improve their competence. At a different level, there are questions about whether central testing and grading procedures help municipalities, teachers, and State agencies make appropriate judgements? The reform efforts, these questions assume, are to make the teacher more competent through processes of participation and better testing procedures of achievement. The questions are embedded in assumptions that relate back to the problem of social engineering as the purpose of state practices; but in a milder, more gentler, and responsive way than in previous eras in we look to other countries.

It are the assumptions of social engineering that, I will argue, need to be made problematic. I want to argue in this section, for example, that the focus on State school "goal-steering" through setting general objectives to be achieved and school-based evaluation in current reforms are strategies that embody profound changes in "the civilizing process"; that is, how individuals "learn" to be productive and "self-motivated" in society. To make this case, I explore certain social transformations occurring in cultural, economic, and political arenas as homologous to the strategies in school reforms and its evaluation. My purpose is didactic -- to use the discussion of homologies to argue in later sections that the educational sciences need to make problematic the concepts that guide their inquiries in light of the changing conditions; that the systems of ideas about social engineering are no longer be adequate to the task of interpreting current practices of schooling.

The changes in the different arenas, I want to argue further, cannot be accepted solely as an argument to include and extend productivity, to provide worker self-fulfillment, or to extend democratic principles in schools. The new citizen, worker and teacher as models of flexible decentralization and durable, transferable competencies are models built on subtle patterns which privilege certain groups and, in Bourdieu's sense, authorize certain social and cultural capital through the dispositions legitimated.

The Post-Modern in the Arts and Sciences: The Aesthetics of Time and Space

In the international arts, architecture, and the epistemologies of the social science, we can identify a changing sensitivity to space and time. "Post-modern" arts, where the term was first coined, accepts a fragmentation of space rather than homogeneity and unity of space (see, e.g., Harvey, 1989; DiMaggio, 1987). The post-modern in architecture, for example, entails different geometric forms and mixtures of styles of architecture whose consequence is to disrupt the relation of time and space which characterize the geometric harmonies of modernism.

The aesthetics of the post-modern arts appears to a modern viewer as a world of fragmented lines, instabilities, pluralities and pragmatic actions. These actions do not seem to fit together into a big story or hang together as one larger set of unifying, universal values that stand as the pinnacle of progress itself. What is important is the local, the plot of the specific without some grand unifying theme or universal message. In the media of television and advertising, for example, there is an increasing view of the world as one of contingency, plurality and poetry of images that seem to have no strong unity/linear sensibilities that have been associated with modern life.

The themes of pluralities and pragmatism are carried, as well, in the epistemologies of post-modern social theories. The well worn phrase that "knowledge is socially constructed" is reiterated in discussions of anthropology, philosophy, political science, psychology and sociology. The belief that science can find a universal knowledge is held by very few social scientists today. The view of knowledge is one that is historically contingent and pragmatic -- knowledge is continually revised and related to, among other things, the social values, social conditions and technologies available in the search to understand. The acceptances of a pragmatic view of knowledge entails an a historical relativism (but not necessarily a philosophical relativism). That is, truth is tied to the complex conditions and methods of inquiry itself. Fluidity, diversity and the seeming break-up of permanence are made into a theoretical precept.

At the same time, the previous reigning Cartesian distinctions between "ideas" and materiality -- mind and body -- are challenged.[4] We can understand the challenge to the divisions between the objective and subjectivity that have guided social and political thought since the late 19th century. One can think of the 1970s social sciences arguments which focused on the "over-socialized person" and the emergence of "qualitative" studies in education as responses to questions about how one can interrelate an objective understanding of structures with strategies to consider subjective dimensions of social life.[5] Current discussions in the social science disciplines, "post-modern" feminism, and neopragmatism in philosophy theorize ways to re-conceptualize the dualism of the objective and subjective; the discussions enable a rethinking of how change is possible and individuality constructed (For example (See, e.g. Bourdieu & Passeron, 1977; Giddens, 1990; Butler, 1993).

With the questioning about knowledge occurring are important shifts in the problems of social study. A historical sociology has emerged to focus on the problems of representation.[6] These studies consider how particular discursive practices construct their subjects; that is, how the categories, distinctions and differentiations of educational theory, for example, constructs what it means to be a teacher and student (Popkewitz and Brennan, in press). In Post-Modern Feminist scholarship, research concerns gender rather than biological distinctions. It considers how women's identities become historically tied to the body and sex, focusing on how women are historically visioned and re-visioned.

I pose these differences in concepts of time, space, and epistemology to initially draw attention to homologies to the reforms and evaluation practices. We can think of the school reforms as themselves re-constituting the social space and identities of teachers.[7] There is an emphasis on the local, pragmatic knowledge that is symbolized in the discourses about school based evaluation and professionalization. These discourses value the practical experiences and autonomy of teachers in the improvement of teaching.

But the valuing of a pragmatic knowledge is not only of the external requirements of the teacher. There is also a different sense of the identity of teachers and administrators who work in schools. The problem of reform is seen as no longer having a final, one right answer but, as one Norwegian administrator argues, it is asking the right question and teachers "feeling of ownership for the knowledge gained and a feeling of responsibility for putting [that knowledge] to use in their own school". Somewhat related to the contingencies of knowledge discussed above, the previous hierarchical relationships of schooling in which a universal Truth was sought are no longer seen as adequate:

It is essential...that we meet as equal, and that we do not possess the truth with a capital T...Our knowledge and insight is based upon a broader experience, and this fact enables us to focus attention and asks questions on particular fields of interest, but usually we do not know the local conditions well enough to, nor should we, give the answer, unless expressly asked to give our opinion (Omholt, 1994; p. 3).

While I do not want to minimize the important distinctions between the arts and social sciences and that are found in the restructuring of schools, the conceptions of personal competence in each arena are built upon and responsive to changing conditions. These changing conditions inscribe new sets of regulation about how the person is to be self-motivated and self-disciplined. In one sense, we can understand the shift in regulation as one from that of social engineering to one where the capabilities of the person's productivity become the site of deployment of change. The different layers of current school reform that range from goal-steering of school to school-based evaluation and reflective teaching embody the new conditions of regulations that I speak about here -- the talk of reform is not about the competence of teachers and students but of the dispositions and sensitivities in which individuals are to feel self-fulfilled, self-motivated, and self-directed. In the following sections on the economics and the political arenas, I focus more closely on the problem of the re-constitution of self-regulation patterns.

Economics and the New Worker:"I understand it" + "I can do it" + "I care about it" = "capacity"

If we examine economic patterns since World War Two, we find major shifts in the patterns of work and the "mentalities" that are to make for a productive worker. Prior to World War Two, the "Fordist compromise" was seen as the formula for economic growth. This entailed a compromise among workers and industrialists in which a labor division and mechanization was accepted with a favorable wage formula and the implementation of a state welfare system. Both the sociology of work and the State projects of welfare were built on notions of social engineering -- that the attitudes, expectations, and knowledges needed for efficiency could be constructed through wise policies and institutional practices. The Fordist compromise, however, no longer works in the economic centers of the world. As the world economy was restructured, hierarchically ordered mass production that trained the worker to be competent in specific tasks of an assembly line lost its efficiency through the production of new technologies and markets (Boyer, 1989).[8]

With a epistemologically ring similar to that found in the arts and social sciences, the principle of new business organizations is "the law of the microcosm" where the more agile, flexible, horizontal structures involve groups of workers with specific projects that do not have the older layers of managements (Fatis, 1992). The smaller units are to "empower" workers and to develop flexible, responsive environments that can respond quickly to customer demands, a language that is not uncommon in contemporary educational rhetoric of reform.

With the rhetoric to "empower" aside, we can sociological understand the new work situations as annulling the arbitrary division between the economic and the social. Work is built on self-managing teams and worker participation in the simultaneously economic and social decisions implicit in the reorganization of production. "Instead of defining the individual by the work he is assigned to, it regards productive activity as the site of deployment of the person's personal skills (Donzelot, 1991, pp.252)".

In a certain sense, there is a homology between the post-modern epistemologies of the social sciences, current theories about the identities of the worker, and educational strategies of reform. Each values a pragmatic, contingent knowledge as productive in one's personal life. The lack of a linearity in the arts, the new worker who is a team worker who problem-solves, and the "identity" structures inscribed in school reform each, but within a different set of relation, construct rules about knowledge and personal productivity that work again the prevailing notion of social engineering as prescribing universal rules to direct action. The emphasis the local management and governing of schools, for example, has included discourses of professionalism. The professional educator to appears in the international literature is a teacher whose competence is not in specific skills and knowledge. For example, Shön (1987), who draws from studies of professionals who work in large organizations, argues the need for individuals to solve immediate problems of their job and have a "knowledge-in action" which requires on-the-spot reflection. This notion of the professional "self" is brought into educational discourses about reform. The professional teacher is constructed as an individual whose capabilities and skills are pragmatically bound to the workplace. They have particular sensitivities and dispositions to enable work in contexts that have fluid boundaries and practically formed problems (For a critique of this view, see, Laursen, 1994).

It is interesting to relate the identity of the teacher as a "professional" to current literature about the identity of the worker in the new contexts of production. The social organization of labor in business and the schools combines technologies with human "factors" (See, e.g, Kantor 1989). The new worker, according to the International Labor Organization report (1994) on the world-wide metal worker industry, is an individual who can work with high levels of quality, productivity, and flexibility. The new "mentality" of the worker is produced as technologies (eg, robots), organization principles (such as "just-in-time" production), and new materials re-vision the production process. The ILO report describes the characteristics of this new worker as a triad:

> "I understand it" + "I can do it" + "I care about it" = "capacity" (p.23)

The conclusions of the ILO report is that the educational system should no longer focus on competencies alone, but on the creation of flexible dispositions towards work; furthermore, that this be accomplished by making the educational system itself flexible and responsive to changes. The report, like many contemporary educational reforms, emphasizes localized program that have a wide variety of teaching and learning methods. Personal characteristics, this report concludes, are central to education, emphasizing self-confidence, self-discipline, the ability to define and resolve problems, and the capacity and willingness to learn" (International Labor Organization, 1994, p. 23).

Politics and the New Citizen:"I understand it" + "I can do it" + "I care about it" = "capacity"

Without going too far afield, but having its internal dynamics that are different from economics, are changes occurring in the governing patterns of state institutions. These changes have two dimensions. The changes have altered the relations of actors in the political decision-making of educating; including the production of "new" actors in the governing patterns. The changes in actors embody, as well, more subtle changes in the competencies associated with the citizen--the skills, attitudes and attributes of the individual who participates in political arenas. But as important, the changes have re-constituted the competence and capabilities of the person who is to act within the context of the schools.

At one level, we can think of the phrases of "decentralization" and "devolution" of power used in current political rhetoric of state practice as more than the shifting of the governing of schooling from a central government to a central/regional coalition. I will argue here that the state re-organizations being witnessed as decentralization and the devolution of state practices are examples of deeper changes through which governmental agencies and civil society interact in the production of social regulation. Various actors within civil society and government have been re-constituted within the field of education. We can understand the various "constituencies" that are given reference in the Norwegian evaluation, for example, as also producing new sets of relations. Governmental agencies are now positioned with a field that includes regional and professional groups to signify changing patterns of regulation that traverse both the political center and local practices. Through examining changes in Swedish political steering of schools, I will argue that we can think about the changes in the field of relations in schooling as inscribing purpose and direction to school practice that are different from what previously existed. The consequence of these changes makes the classification of state and civil society as no longer meaningful.

If I think about socialization theories of the "citizen" in the 1970s, they focused on a person's skills and competencies to act in political arena. A universal set of skills and knowledges for citizenship was assumed. American liberal political theory of the 1970s, for example, defined the citizen as having particular characteristics of participation that were to be developed if democratic political institutions were to function. Citizens were to belong to multiple civil associations, behave in a particular rational manner as their participation sought to influence governmental policy. Schools were to teach about the legal system of government and provide opportunities for students to learn the "rules" of participation, especially the importance of being literate to assess the qualities of candidates and to vote wisely in elections. It seemed clear who the Democrats and Republicans were.

The efforts to make the socialization pattern for citizenship more efficient and effective was a classical case of social engineering. The 1960s U.S. War on Poverty, for example, was designed to remove the obstacles of participation among the poor and thereby, it was believed, to eliminate the psychologically debilitating effects of poverty.

In the current situation, such clear lines about participation and competence are no longer available. If I think about the U.S. and Europe, political parties agendas can no longer be given reference to through notions of class as in the past in Europe. Ideological positioning of policies as conservative, liberal, and left do not provide explanatory concepts to understand how political parties operate. Social movements and political projects of the past decades, further, have made the boundaries in which the "citizen" acts less visible and distinct. New expressions of political projects are more local and less class-focused, such as in the "anti-system" Green movements and the politics of feminism.[9] The social, "anti-system" movements place increased emphasis on the politics of "identity" and on "grass-roots" political ideologies rather than on party politics. The identity of the citizen in the new contexts of the "political" gives focus to the individual's capacity not competency: one who "understands", "does", "cares about", and acts with greater autonomy, flexibility, and problem-solving ability.

We can explore the re-constituting of the identity of the "citizen" through examining the changing patterns of state governing in Swedish schools (Popkewitz, 1993c). Changes between central and local governments are not only who makes decisions. The changes are in the governing patterns through which the professional and parents in the school are to act and talk about teaching and learning. Previously, the Swedish school was organized through rule-governed legislation. Such organization entailed detailed parliamentary instructions to teachers that were to ensure policy implementation in the details of classroom instruction. The responsibilities of organizing school subjects belonged to the state ministry and bureaucracies, not the teacher. We can understand state policies to "rule-govern" as part of the mentality in which the state was to socially engineer progress.

New governing practices, in contrast, involve a particular pragmatic outlook which is goal-governing rather than rule-governing. The recent formation of the Swedish Agency For Education (Skolverket) and reformulations of Swedish curriculum (Läroplan) involve the implementation of this new goal-driven conception of the state *vis-a-vis* the educational arena. General goals are set by the state, and to act as a "steering" mechanism by which local communities identify solutions appropriate to their setting. A problem-solving attitude is hereby constructed which accepts a plurality of solutions to social problems.[10]

The new governing patterns of the state have also introduced new governing patterns of the teacher and the state bureaucrat who is to monitor educational practices. Discourses of professionalization, didactics, curriculum and teacher education place an emphasize on the particular types of sensitivities and dispositions of teachers. If I draw from research in Finland, Sweden, Iceland, Portugal and the US., for example, reform has stressed teachers' practices of "action research" and of being "reflective", but these practices need to be understood as more than simply making teachers more competent and professional (Popkewitz, 1993a). The practices are embodied in the construction of State steering of schooling through goals rather than rules.

In the new context of governing, teachers' competencies are tied to being able to be responsive to goals that are defined generally rather than specifically, and to deal with the contingencies of local situations in which the outcomes are to be implemented and judged. The shift in the political organization of the school and teaching also introduces a State official who must now monitor school practices in a manner that allows for multiple solutions to problems and who cannot expect prescribed answers in State goals and procedures. This "mentality" of state officials is very different from that in a rule-governing system.

The changes in political steering do not stand alone, but are inscribed in an amalgamation of practices whose consequence is to alter how teachers judge their competency and order their reflection. Simola (1993), for example, has argued that Finnish teacher education has shifted from a seminary and missionary tradition where moral/religious cosmologies and "craft" tradition were valued. The previous "craft" traditions can be understood as tying performance to an apprenticeship model of learning; and where moral purpose explained and interpreted teachers' actions. But we should not think of the seminary model in relation to the Middle Age notion of "guilds and crafts," but to a 19th century conception that was chanage-oriented and associated with the spread of modernity. In the United States, the teacher was to maintain the Protestant images of pastorial life while being an agend of change. The new discourses of teacher reform are related to practices whose norms of competence are tied to scientific rationalities (Novóa, 1993; Simola, 1993; Johansson, 1993; Kallós & Selander, 1993; Popkewitz, 1991).

The shifts in reasoning about teaching are not only about norms of competence but also about who is authorized to speak about teaching and teacher education. The newer scientific cosmologies in which the reform concepts of "reflective teaching" and "action research" are placed, for example, give value to a professional expertise that seems, at one level, related to a practical reasoning tied to teacher wisdom and experience. The "teacher" in the reform literatures is a professional whose practical experiences provides an expertise to organize instruction. But that "practical expertise" is not of the seminary tradition nor of the "words" and ideas of teachers. The "practice wisdom" of teachers is organized through scientific systems of classification and which are applied to assess individual teacher competence.

The "new" disciplinary principles supersede seminary traditions and reconstruct the principles through which teacher competence and capabilities are to be assessed. In this context of changes rules of competence, we can understand the re-emergence of studies of the didactics of school subjects, for example. This field of research rationalizes and provides criteria for judging pedagogical practices. In the United States, the didactic expertise has been called "constructivist pedagogy". The discourses of didactics emphasize a teacher who is pragmatic and problem-solving, characteristics that are homologous to the principles of work, state steering, and the epistemologies of the social sciences discussed earlier.

BEST COPY AVAILABLE

Homologies and The Construction of the Teacher

We can bring together the previous discussion as a way to form an horizon in which to think about reform discourses of professionalism, centralization/decentralization, school-based evaluations, as well as to the conceptions of participation that occur in educational reform. These changes involve two sets of regulation. One was regulations produced through a field of actors who are authorized to classify and speak as the legitimate spokespeople about schooling. And second, and often less considered is that of "the civilizing processes; that is, the regulation that occurred through the categories and distinctions that construct the person who is to act. In the previous discussion, for example, I focused on multiple arenas to consider shifts from the individual as defined as particular sets of competencies to the individual who embodied pragmatic capabilities and dispositions.[11]

The structuring of "mentalities" is not explicit but occurs as aesthetics, epistemologies and work conditions interact and produce ways of "seeing", thinking, acting, and talking about the world and individuality. Reforms, for example, have conceptualized school knowledge as contingent and plural. A focus is placed on localized programs and teacher professionalism which, in turn, demand flexibility and responsiveness to the changes. Further, the "capabilities" of the teacher are "self-confidence, self-discipline, problem-solving, and willingness to learn. My repeating of the equation ("I understand it" + "I can do it" + "I care about it" = "capacity") in the sections on economy and politics was to suggest homologies in the construction of the "subject" that relates to educational reforms and evaluation.

My thinking about the changing "civiling processes" is to pose the problem of change as relational, that is, to consider different arenas as intersecting and weaving. The intersection of the different arenas are inscribed in educational discourses about professionalization and school-based governance which value individual "self-confidence, self-discipline, the ability to define and resolve problems, and the capacity and willingness to learn". This "seeing" of things relationally, however, must recognize that a historical specificity as well. There are differences among the arenas that Berger et al (1973) argued when they considered the differences in the structuring of thought, belief and cognition within industrial production and state bureaucracies.

Further, the consequences of the new changing patterns of regulation are not unproblematic. As I mentioned above, the changes inscribe not only different discourses about the capabilities of teachers, the changing patterns also inscribe certain types of expertise as the sciences of didactics become the authoritative voice in classifying the performances of teaching. The shift to scientific discourse also have raised issues of gender. In research that I and Kathryn Lind (1989) have done concerning reforms to the professionalization of teachers, for example, the professionalization practices entailed teachers constructing assessments to more systematically gauge the consequences of their practices. The assessments inscribed a particular rationalization of teaching whose consequence was to devalue the cooperative, "caring relations", and child-centered pedagogues given priority in the classroom by many of the women teachers.

The "looking" at the different arenas require, I believe, that the methods of evaluation that question the changing boundaries of teaching and learning occurring. That is, when we talk about teachers becoming involved in school decision-making or a decentralized system of governance that will promote greater liberty, social justice, and professional autonomy, we need to recognize that the words exist in sets of social relations that need to be made continually problematic. It is this problem that I turn to in the remainder of the paper.

THE DENIAL OF CHANGE IN THE PROCESS OF CHANGE: OLD CATEGORIES AND NEW CONDITIONS

If we consider the previous argument that related school reform and evaluation to other social arenas, we find important changes in the identities of teachers and students that are being constructed. The categories of evaluation, I want to argue in this section, however, often deny those changes. The denial occurs in a least two different ways. First, is the seductive quality in the rhetoric of school reform and evaluation. I use the U.S. reform rhetoric to illustrate how a language of localism, professionalism, and local school management produces images that make current practices seem as part of a long term effort to realize liberal democratic commitments. But the language is seductive, its obscures how current condition have different assumptions and implications than what existed previously.

A second layer of rhetoric concerns how concepts of evaluations are caught between images of the past and conditions of the present. I will argue that current assumptions about the state and civil society (carried in the categorization of centralization/decentralization), "frame factor theory" to discuss restraints on teaching, and the concepts of participation in the evaluation are frameworks that emerge from past conditions of a presumed socially engineered society; a conception of change that I argued in the previous section is no longer appropriate.

The Seductive Reasoning of Reform: Populism and Prophesies

We can understand educational reforms and evaluation as rhetorical styles that are seductive (Rooney, 1989). They are seductive through the establishment of discourses which make the "new" conditions seem as merely a continuation of past social\political commitments. I focus on the rhetorical strategies of the U.S. "systemic school reform" movement to illustrate the seduction; producing images that have some resonances in the Norwegian discussion. There is an attempt to find a proper balance of state coordination of school practices with local and professional practices. At the same time, the U.S. rhetoric, as like the Norwegian, brings into focus the social and political commitments towards an equal and justice society. My consideration of "systemic school reform", then, raises the issue of how rhetorical styles can obscure that conditions which are to be evaluated. In the following section, I focus on how the concepts of evaluation place past conditions on those of the present.

The premise of "systemic school reform" is seemingly simple: there are a myriad of reform practices to improve schools as the collective wisdom of research has enabled us to identify the characteristics of "successful" schools.[12] Governmental practices, however, have not been able to capitalize on this knowledge because reform efforts have been too fragmented, uncoordinated, and incoherent for that knowledge to be useful on a large scale.

The rhetorical form of the argument is seductive. A populist, progressive language is introduced: A principle of the reform "is the notion that if school personnel are held accountable for producing change and meeting outcomes objectives, they will expend both their professional knowledge and their creative energies to finding the most effective ways possible to do so, relevant to the specific conditions in which they work". Reference is continually made to a need for "a fundamental rethinking and restructuring" of the school, calling forth the phrases of "decentralization, professionalization and bottom-up changes" as "keys concepts" (Smith & O'Day, 1990,p.234).[13] The teacher and child is to be "empowered!"

One can read the phrasing of the problem and solution as re-iterating general and long-term populist commitments: the people in "bottom" positions, teachers, are most knowledgeable about what needs to be done to improve schooling, and therefore school personnel are in the better position to make long lasting change. The language of reform also ties its practice to historical social welfare commitments in using the U.S. schools to produce a just, equitable, and intellectually challenged citizenry.[14]

The populist, progressive thought of reform joins a particular American "civil religious" discourse of prophesy and redemption. The need for systemic reform involves a warning that "unless coherence and clarity" is given in school policy and practice, the "relative quality of the education offered to less advantaged students" will be eroded (p.262).

> Simple justice dictates that skills and knowledge deemed *necessary* for basic citizenship and economic opportunity be available to *all* future citizens - that is, access must be distributed equality, not just equitably. (p.263)

The words provide a sense of continuity and continual movement towards a progressive realization of the goals of schools. The phrases could have been said in the U.S. during the 1960s, 1970s as well as now.

The sense of commitment and of rhetoric is no less a practice in the Norwegian educational arena, although there is a different historical tradition of the State and in the role of science and professional expertise. The discussion of school-based practices, teacher reflection and action research occur as a normative concern about the importance of local autonomy and development of schools; and the need to examine changes in curriculum in light of principles of equity and justice. The U.S. concepts of "decentralization", "professionalization", and "bottom-up" reform resonate with those of the Norwegian discussions about schooling and evaluation. Even the discussion of "frame factors", a more general Scandinavian concept about the influences on school practices, continually returns to a joining of local and individual responsibility with State responsibility to ensure equality of opportunities, access, and human rights- the very problem that is set within systemic school reform.

While I recognize the need for prophesies and the importance of populism, I also recognize that rhetorical strategies of the reform are not only seductive. In "systemic school reform", the meanings of particular words should not be read as only a reiteration of past commitments but as words whose meaning are produced in practices that are not necessarily spoken in the text. In part, these practices relate to the problems of regulation that I focused on earlier in the discussion of the arts, economy, science and politics.

While populism has had moments with radical implications within the U.S. (Goodwyn, 1978), we need to recognize that modern prophesies about a better world are built on experts in the services of the democratic ideal. There is a long historical tradition of the social sciences as tied to state planning and evaluation of the projects of modernity (Ross, 1991; Wagner et al, 1991; Popkewitz, 1984). If we examine the discourses about "the professional autonomy of the teacher " and "self-governing" that criss-cross in the "systemic reform" movement, for example, we realize that phrases are not neutral but part of social relations that give reference to the employed words. "Self-governing" is not the teacher who acts with an independent "voice". The teacher is one who is disciplined through the new state patterns of coordinating and setting standards. It is also a teacher who participates in a professionalized hierarchy in which traditionally inservice and preservice programs are conducted. The rhetorical strategies of "bottom-up" reforms and governmental practices of coordinating and giving coherence assume a neutrality in a context that has changing conditions of power, a problem I return to again in the final sections.

Old Distinctions and New Conditions?

My task is now to understand how the words of the evaluation do not stand alone but are placed within social fields through which the words are given meaning. To do this, I focus on the distinction between the state and civil society that I explored initially in the discussion of changes in Sweden and which reappears in the reform and evaluation discourses. The distinction permeates our systems of ideas in ways that we are not aware of; carried in words such as "decentralization", and "frame factors" that restrict school practices, but also in the ideas of "professionalization" that are historically part of the construction of modern state patterns of regulation and social engineering (see, e.g. Ross, 1991; Popkewitz,1993b). The categories of "regional decision-making" and "local autonomy" in the evaluation, as well, impose the distinction between the state and civil society.[15]

The distinction, however, is not only conceptual -- it is normative. In the United States and Scandinavia, the State is viewed as separate from civil society but as a means to intervene in civil society to promote social goals. One intent of a State evaluation, I think without going too far afield, is for social engineering -- to further democratic processes through its efforts to increase professional participation, to facilitate the development of governing in local communities, and to foster commitments to foster equality and justice in the practices of the society -- but again, the strategies are "softer" and constructed in a manner that has "post-modern" as well as modernist tendencies.[16]

I raise the distinction of state civil society to focus on how this distinction is used in contexts in which it is no longer analytically descriptive. While I argued that the problems of social engineering have become "softer" and more pragmatic in the changing conditions of schooling, the working conceptualizations in evaluation practices in school-based evaluation and measurement often inscribe a dualist relation of state and civil society that are being problematized in current conditions.

We can think of the dualism of state/civil society as theoretical "entities" of the late nineteenth and early twentieth centuries which introduced separate spheres of action that could be planned, supervised and evaluated in the modernizing project of the State. Implicit is an Hegelian assumption of the State as the embodiment of higher reason which mediates among different social interests to produce progressive projects. The concepts were part of a map that separated the social world into different arenas, such as the private (home)/public, economic/social, and state/civil society. What I want to explore here is how our use of such words as "democracy", "community" and "participation" in current policy discussions of schooling maintain the dualism of the categorization of the "political" as embedded in state/civil society and its potential limitations for understanding the consequences of educational practices.

We find that if we examine the social practices of the early 20th century, the dualism did not work even then. The processes of governing merged the political, social, cultural and economic rather than separated them. The emergence of the welfare state, for example, joined the public/private as the macro State agendas of administering liberty were linked with the micro construction of the citizen, what Foucault (1979) called "governmentality". U.S. Government, for example, extended its activities in the first half of this century in a manner that rapidly encroached upon the private lives of citizens: the state guaranteed people an income after retirement, paid people who were out of work, subsidized corporate and agricultural enterprises, and created jobs for people, and began to subsidize the development of psychotherapic practices for working through social issues as personal problems.

The discourses of the social sciences were a central strategy in this process of joining the macro problems of governing with those of the lives of individuals. The political, social, and economic upheavals in Europe and the U.S. during the first decades of this century produced a series of social theories that responded and interpreted the crises. Merelman (1976) argues, for example, that the emergence in the U.S. of a behavioral political science was can understood with the changing social conditions. The political theories were to enable people to cope with the emotional gap between things as they were and things as people would have them. Political theory constructed symbols that reestablished the ideals of American politics in the changed political circumstances. The symbols included notions of a "political culture", "pluralism", and "political socialization"-- concepts that resonated with pre-existing beliefs about political community and public consensus but tied these beliefs to the realignments of governing occurring in the 1930s.

The political theories were symbolic canopies through which the subjectivity of the "citizen" could be re-visioned. The new techniques of political inquiry of polls and surveys of public opinion, for example, enabled people to believe that they were being consulted and that it was possible for people to act purposefully in the changing circumstances of governing. The theories and methods of behavioral political science, Merelman argues, helped reduce strain between the changing role of government and the cherished beliefs about public life.

I pursue this history of American political science as instructive to my point of understanding how the governing processes are changing and how past distinctions can obscure those changes. In the U.S., the theories of politics made the distinction between state and civil society; but the idea of participation entailed shifting categories and distinctions through which the person was to constructed as a competence citizen. The theories of political participation redefined the way power was effected in daily life but hid the effects of power through a language that portrayed the distinction of state and civil society. It hid the effects through obscuring the ways in which macro politics were brought into the very conceptions of the "citizen".

It also concealed through constructions of concepts of participation that were defined as universal rather than multiple, related to particular patterns of power, and differential in assumptions and implications. The recent work in feminist studies of political theory, for example, illustrates the inclusionary/exclusionary qualities of models of participation. The 19th century notions of participation to spread democracy were constructed with a particular gender bias -- describing the attributes for participation that, first, separated the public and the private and then defined a public rationality in a manner that was to exclude women (Pateman, 1988; Lloyd, 1984). The "nature" of participation privileged the manners, views of rationality, and public discourses associated with the "reasoning" of men and thought in opposition to dispositions that women had in the privacy of the home.

In the current situation, we find a harkening back to old categories that do not interrogate the different sets of relations and different constitutions of capabilities and competencies associated with participation. The change in relations and capabilities, as I argued earlier, can be understood in Sweden which has major changes in the past decade in its governing processes of education. These changes are not only in the organization of education. They are in the discourse that constructs, legitimates, and normalizes questions and practices associated with education -- practices that were previously outside the range of what was permissible to question, such as the exclusions of certain ethnic groups from participation, or a re-examination of the "social democratic model" that relocates the beginnings of welfare state in the 1920s, prior to the election of the Social Democratic Party.

Swedish analyses of educational policy continually point to historical disjunctures in the discourses about citizenship, of the rupture between state policy, bureaucratic discourses and the cultural perspectives that are to define the child and the teacher in the school. Analyses of the reforms, for example, register shifts in the values that order the early childhood and the comprehensive school (See.e.g.Lindblad & Wallin, 1993; Englund, 1994a; 1994b; Dahlberg & Åsén, 1994; Walsh, no date; Carlgren, In press). Significant moral and political questions in the comprehensive school and early childhood, it is argued, have been refocused to technical questions of economic relations and management. The changes are identified as a movement from a previous social democratic model (that emphasized Fabian socialism, social democracy and social engineering) to a mixed model that steers the school through market forces and privatization, and through state intervention to build a common school. The "new" Swedish school is identified as having multiple definitions of pluralism and democracy; such as those related to alternative school movements, religious fundamentalism, and different profile schools for parents to choose for their children's education.

While the literature points to current policies as a rupture from past social welfare commitments, the dualism of the state/civil society is reinscribed through a sense of history as a process that is either evolutionary or of a cycle with the State as the governmental whose "hand" guides historical changes. The words to interpret the Swedish reforms, for example, construct particular sets of images of the past in the present; such as a "New Right", "restoration tendency", "reestablishment of traditional schooling", a "return to early curriculum traditions", a "weaker curriculum steering", and "step back to traditional curriculum".

The central interpretive focus is the State government as organized by the party in power rather than the state as an epistemological concept to explore the historical set of relations around the theoretical issues of governing are interpreted. The critiques accept the normative assumptions of what Norwegian sociologists have called the "holy triangle" of power; that is, the alliance of the governmental, The Social Democratic Party, and researchers in the production of social engineering policies.[17] The politics of the Swedish Conservative government who replaced The Social Democrats in the early 1990s are juxtaposed in educational literature with that of Britain's "Thatcherism" and the U.S. "Reaganism". The Conservative political language of "market" and "privatization" are re-inscribed in research discourses to assert the dangers of policies that challenge to the Holy Triangle.

While it is clear that the moral and political terrain of educational struggles have shifted, such analyses beg the question of the historical conditions which make these words (and policies) plausible among different national contexts and parties. Again, if the previous section about politics, arts, science and economics are historically appropriate, the changes that we now witness in the school arena are homologous to changes that have been of a long duration in multiple arenas, long before Reagan and Thatcher took office. *The struggles about "markets" are also part of multiple struggles about citizenship, childhood, work, gender, and schooling that intersect in this particular historical conjuncture.* To focus on particular political parties loses site of those changes and their implications and consequences. But the analysis also creates a conundrum as conservatives, liberals, and social democrats engage in similar policies (such as the educational reforms of the Socialist Party of Spain and the U.S. Democratic Presidents of Carter and Clinton).[18]

The need for changing interpretive schemes in which to interpret and evaluate the current is brought to the fore in Carlgren's (In press) analysis of the Swedish Committee on the Curriculum in 1991. As a member of a Swedish Committee on the Curriculum in 1991, she sought to historicize her position through understanding the complex social relations in which knowledge of the curriculum was constructed. The analysis places governmental practices within a complex network in which the terrain in which the struggles about curriculum ideas can be interpreted. The public debates about the new steering documents for the new curriculum in 1990, Carlgren argues, were framed around old categories of schooling (changes in the school timetable) and, as a result, the public discourses missed the ways in which the steering by objectives and results re-worked and imposed new rules for what is said and written about. She argues, for example, that the curriculum guidelines shifted from a process language to one of production language, subject requirements and assessments. The curriculum re-visioned of the knowledge of schools and has consequences for what purposes of schooling can be realized.

In the analysis, it is not the state/civil society distinctions that centers the interpretation; but the ways in which various "actors" and discourses constitute governing patterns in the school arena. The formation of the curriculum goals embodies a complex networks of actors from which the problem solving about schooling was constructed. In the political arena discussed earlier, for example, we saw the various practices of the Swedish agency Skolverket as themselves embedded in changing patterns through which power was deployed. To speak of the organization of power as that of the political center or of local school governance obscures how the patterns of power are being re-constituted (also see, e.g., Weiler, 1990). Further, to define the problem as "New Right" and Neo-liberal political practices removes from scrutiny how different groups are positioned in the practices of governing, including researchers and professionals.

Participation needs to be conceptualized within these relations of actors and discourses. Further, and as important, the practices of school-based evaluations, reforms such as action research, and the new measurements systems through which to judge school competence do not stand alone. They are inscribed and to be interpreted in the amalgamation of relations that are being constituted.

My argument to historicize the concepts that are embedded in the research and evaluation communities has been to make the "givenness" of the concepts of schooling, reform and evaluation as part of the "things" that an evaluation needs to scrutinize. If the previous discussion of changing patterns of actors and identity were appropriate, the evaluation practices are located in social configurations that are qualitatively as well as quantitatively different from those of earlier in the century. The construction of time and space in art, the worker who is has the flexible capabilities of a "problem-solving" in the metal trades, or the professional competence of the teacher in a goal-steering process described earlier, inscribe certain normativities in how individuality is to be understood and act upon. These normativities cannot be understood through the state/civil society distinction that is embedded in conceptions of participation and professionalization, nor the meanings of citizenship or childhood; rather researchers and evaluators must understand how these terms are placed, meaning derived, and power constructed in the historically contingent sets of relations in which the reforms are practiced.

In the following sections, I want to examine more closely the ecology of reform and evaluation by considering two dimensions in the doublets of inclusion/exclusion. I again use the exemplar of the "systemic reform" movement in the U.S. in this discussion. My purpose is to further place systems of ideas as systems of action in power relations.

EVALUATION AS INCLUSIONS/EXCLUSIONS

The public and research attention in evaluation can be understood as inclusive strategies of reform that derive from long term commitments of the Enlightenment and modernity. Those commitments are the tying of reason and rationality to produce social progress *and* individual freedom. In the school arena, we can understand reform strategies as giving attention to these commitments through the discussions of local, community decision-making, the role of the state vis-a-vis the regional authorities, and of professionalization strategies to provide teacher with more local autonomy through "action research" and "reflective teaching".

But the joining of reason and rationality has produced historical tensions that are embodied in national evaluations; to open up and manage spaces for individuals to act with greater freedom is part of a doublet; the strategies produce boundaries that exclude some from that opened space. The doublet is analogous to map-making that constructs "imagined communities" as the categories of the map that mark off the territory and delimit who and what is to be included. If I turn to 19th century European map and census data (Anderson, 1991), for example, I find "imagined communities" were constructed to illustrate the separation of the colonialist and the "natives". The map-making "told" who "belonged" to the nation. With that belonging was a reclassification of the "native" as different from the colonialist, and, finally, categories were constructed about histories of what was to be remembered and forgotten.

If we think about school reforms as "maps", we recognize that certain preferences, ideas and behaviors of particular groups are made into universal attributes of change; thus, de-legimating or omitting from public scrutiny other possibilities.[19] In this section, I pursue two problems in the "maps" of school reform and evaluation. One relates to the strategies of problem-solving to improve school teaching. Here, I again return to the U.S. "systemic reform movement" to illustrate how the discourses to improve teaching and learning are doublets of inclusion/exclusion. Second, I consider the alchemies of the school subject to raise questions about the epistemological assumptions of measurement of achievement in national evaluations. If we consider evaluation as part of the maps being drawn, then a major concern of the evaluation is a double self-reflectivity -- that is, to make the question of how the constructing systems for administering the spaces for schooling also excluded others from that space.

The Normalizing about Individual Competence

A strategy to consider the doublet of inclusion/exclusion is to examine the problem-solving methodology embodied in school reforms. I return to the "systemic school reform" movement and its approach to didactics called "constructivism." (In Sweden, this field of research is soetimes called "progressivists" and is situated in in a different set of relations among researchers and teacher educators than in the U.S.) Constructivism is a psychological discourse about teaching (didactics) that, like other elements of the educational reforms, are to open spaces for teachers and students to have more autonomy and responsibility.[20] The reforms focus on individual's "capacities" and "dispositions" which will enable individuals to become "self" motivated and self-directed (See, e.g., Interstate New Teacher Assessment and Support Consortium, 1992). Attention is given to the subjectivities of teachers and students; they are to have beliefs and attitudes that "empower" them, make them flexible problem-solvers, and learn in cooperative groups.

The "constructivist" teacher can be understood as having capacities that lend themselves to the effecting the governing presupposed in the organization of school-based evaluation, teacher professionalism, and the goals-steering governing practice. The constructivist teacher is self-motivated and problem-solving as teachers and students "make knowledge and meaning" to involve children as "active individuals who construct, modify and integrate ideas" (The National Council for Teachers of Mathematics, 1989, p.-7). The knowledge of constructivism is a "doing" in which, more often than not, classrooms are considered as a negotiated order: children are to learn subject matter through processes of learning that involve multiple strategies and meaning. Teaching and teacher education reforms related to "action research" and "the reflective teacher", which are part of the Norwegian evaluation as well, carry a similar orientation towards the motivation, skills and capabilities of the teacher.

One could argue that the reform strategies are reasonable attempts to open up classroom learning to greater interaction among children and their teachers,to provide a more nuanced and critical assessment of what is done in teaching. But my argument is not against this assessment; there is ample evidence in the U.S. schools that such approaches can provide for better learning. My argument is that there is a continual need to look at what seems benign and inclusionary are also practices that exclude.

If we consider constructivism as an example of a specific technologies of reform that embody rules for problem-solving, we can inquire how the practical implementation of procedures in an evaluation embody the doublet of inclusion exclusion, such as in the use of constructivist approaches to evaluate learning or the testing approaches to measure achievement. Viewing the specific evaluation technologies as a style of problem-solving enables us to ask questions about its rules and standards, such as what and who are included and exclude. As one can see at this moment, my thinking of the rules of problem-solving is sociological rather than a psychological. For example, we ordinarily think of problem-solving as what we do to find an answer - what is the algorithm applied in mathematics; or what procedures do we use to construct interviews. But we tend not to recognize that answering a problem (problem-solving) requires certain prior conditions and rules that direct attention to what are are not problems and how solutions are to be effected.

The example of constructivist approaches to didactics is illustrative when we consider that its rules of thinking are presumed as universal and applicable all children. The problem is how to provide efficient lessons so all children can solve problems in flexible ways. The "map" about reasoning seems to exist without any time or space dimensions. All distinctions among groups are made into a sameness and there are no other seeming possible maps in which children think or act.

Yet, the reform assumption of a universal subject who "makes" knowledge ignores how social distinctions are made into distinctions of "cognition" and practice. An oppositional space is constructed as constructivism names the children who need remediation or special assistance while, at the same time, asserts a universalism to its systems for classifying how thinking occurs. An example of the universalizing occurred recently as a colleague was writing about the teaching of mathematics to Hispanic students in Los Angeles.[21] The teacher did all the things that the constructivist approaches argued against. He stood in front of the room and lectured. He worked from the textbook, and did not involve children in "cooperative learning". But the students' achievement did skyrocket as did their appreciation of the beauty of mathematics. When the example was given of good teaching, my colleague was told that it was not good teaching because the children did not work in small groups and use manipulative materials.

A doublet of reasoning was inscribed in the pedagogical reasoning: there is the "good" student who acts according to the universal rules of constructivist pedagogues, and an unspoken norm about a "bad" student who does not follow the rules of working in groups that are inscribed in constructive pedagogies. The norms of "good" are not explicitly stated but normalized through the universalization of what is school problem-solving. The problem-solving child stands in opposition to the child whose ways of acting, thinking, and speaking embody different cultural dispositions and sensitivities than embedded in the universal rules ascribed by the constructivism.

Why is the pedagogical political correctness of constructivism socially important? We think about the above example as one which is not unrelated to the construction of evaluation strategies to assess the teaching of school subjects. The exclusions of the pedagogical reasoning are not overt but inscribed in ascription of a universality of reasoning. A particular social groups' reasoning is sanctioned as universal; the principles of the universal exclude "other" groups through a socially positioning that makes certain norms, dispositions and awarenesses, and behaviors as legitimate (see, e.g., Bourdieu, Hertfeld, 1992, Zerubavel, 1993; also see, e.g, Ladson-Billing & Tate, 1994; Ladson-Billings, 1994).

The universalizing of reason that is inscribed in school reforms and evaluation is of consequences when we think about how different social groups are served through the processes of schooling itself. In continental Europe, the processes of such inclusions and exclusions have been discussed as the **2/3rds solution** (Wagner, 1994). Social reforms in Europe, Wagner argues, have emphasized a conception of the individual as an "enterprising" person, whose characteristics are not different from those previously discussed in the arts, politics, economics, and as presupposed in school reforms. The person is assumed to be self-reliant, goal oriented, active, and reward-focused. But the "enterprise culture", Wagner (1994) argues, was a conception for the two-thirds of the society who have the available sentiments and dispositions to create "opportunities". The "Other" one-third was excluded from "the main spheres of society in which social identities can be formed." The "others" may be people of different ethnic and racial backgrounds, or of class and gender as they overlap in the constructions of oppositions groups. What is important is that the exclusions do not appear as "prejudice" but as the setting of norms and standards whose consequences are to make certain social distinctions as legitimate arbitrators of success.

At this point, one might suggest that school curriculum is a normative discourse and there is an obligation to assert certain types of problem-solving as worthwhile and appropriate. The didactic principles, it can be further argued, drawn from constructivist psychologies, is just one example of taking normative purposes of schooling and placing them into the professional obligations towards the improving of instruction. My argument, however, does not reject the normative qualities of school curriculum and teaching, or the general responsibilities of schools to regulate the types of problem-solving. My argument, in contrast, is to explicate the need for a continual vigilance in evaluations through examining the doublet of inclusion/exclusion. As with my earlier discussion of participation, the universal conception of problem-solving ignores how particular sets of social distinctions are socially constructed. As Dumm (1993) argues in a different context, the discourses of the social sciences are normalizing practices that classify marginalized groups such as people of color as different from the norm and who, *at best*, can be "like the normal person".

Thus, the different school-based strategies to professionalize teachers and the changing definitions of didactics which pervade the evaluation must include questions that not only ask what forms of participation is included, but also who are excluded. While the social spaces for school practice may be enlarged, the "capabilities", motivations and interests privileged in the evaluation are also practices of exclusion.

The Alchemies of School Subjects

In this section, I want to move the discussion of inclusion/exclusion to the knowledge in school subjects presumed in reform programs. The testing procedures instituted in national evaluations in Sweden, the U.S, and Norway, among others, assume that what they are testing what is described by its label; their test items in fact test literacy, science, mathematics, social science, arts, etc. I want to argue that these evaluation practices take-for-granted the knowledge of the school subjects, not recognizing the alchemies that occur in schooling. As with the sorcerers of the Middle Ages, school subjects are alchemized as there is a passage from the social spaces of disciplines (such as the sciences of physics) into the social space of schooling-- that is, there is a movement from the production of knowledge in the "culture" of physics to a school curriculum knowledge called physics (see, e.g. Bourdieu, 1990 who discusses alchemies in relation to the social sciences). The movement from the field of physics to the field of schooling presupposes a change in mental ordering of knowledge. Physics becomes a problem of school learning but that change is not acknowledged in pedagogical theory.

An effect of the school alchemy is "imaginary subjects" and "imaginary practices" (Bernstein, 1992). The forms given to learning school subjects have no basis in the world outside of the school. The distinctions about science or mathematics conform to the rules of schooling that have little to do with the original disciplinary fields. What is brought into school is not what scientists, mathematicians, writers, or artists do; rather the disciplinary knowledge is reformulated to conform to expectations related to the school timetable, to conceptions of childhood, as well to conventions of teaching which cast such knowledge into a school curriculum.

The social space of school is occupied as a world of social psychology, psychology and group management. Science is teaching "cooperative small groups", or "whole-group instruction," or being "caring", "developing self-esteem", "being able to apply the concepts in real life situations". It is a children's world and a teaching world that is expressed as conversations about "helping others learn," "identifying a concept", "evaluating students' assignments", or "assessing student's progress", and "managing the classroom". The net result is something different from what is borrowed.

School ceremonies, ritual performances and discourses, however, make it seem as though one is, in fact, "doing" science or social studies. Sitting in a science laboratory with all its paraphernalia, writing mathematical formulas on a chalkboard, or reading textbooks labeled as physics, rhetorically assign a relation between disciplinary knowledge and school knowledge. The school textbooks, for example, make symbolic links between what is done in schools and disciplinary fields outside of schooling-- textbooks "tell" about the concepts and information of science, mathematics, and social studies. In a high school mathematics classroom in which I was doing research, the teacher told the children that they should act as "mathematicians", talked to them about the importance of learning, and then proceeded to have them prepare for an examination through doing textbook exercises.

I raise the issue of the alchemy because of the centrality of testing and measurement of school subjects - literacy, science, mathematics -- occurring in the evaluation. The evaluation needs to consider the school subjects as not givens to be measured, but practices themselves that need to be inquired into as to the assumptions about what is taken as science, mathematics, social studies, and the arts, among other school subjects.

School subjects are treated as thought no alchemy exists especially when considering the testing and measurement of school knowledge. While the measurement procedures and technologies in school evaluation have become technically sophisticated, the conceptualization of school knowledge is that of the alchemist. That is, if we consider words of measurement such as "concept attainment" and more recent formulations of learning as "conceptions/misconceptions" of scientific concepts, we can recognize examples of how subject knowledge is viewed as *things of logic.* In each instance, the concepts of curriculum as viewed as having essential definitions from which to assess children's understanding. The concepts, generalizations and principles of school subjects are treated as logical and analytical "things" to be learned.

If I go back to an earlier discussion of the social sciences, I can contrasted the view of school knowledge as "things of logic" with a "messier" and more pragmatic understanding of the world which can be called *the logic of things* -- scientific knowledge embodies a variety of complicated and contingent practices as people interpret their the world and, as such, the understanding of science, mathematics, arts, and so on are always more than can be expressed in any purely logical or analytic system. Scientific concepts, for example, are pragmatic understandings that weave together a complex network of methods, social relations and social values that enter into the equation of what is known. Further, the invention and innovation in the sciences entail strategies to make the known-as-unknown and to raise questions about what is taken for granted; just the opposite of school curriculum which privilege the logical and stable properties of knowledge.

Yet if we consider the problems of measurement of school knowledge, we find it is the "things of logic" that dominate the measurement practices of instruction. Achievement is measured by how well subject content is learned. Teaching about concepts and methods occurs "as if" they existed as uncontested and fixed objects that represent some reality. The operationalizing of the appropriate definitions to tests (such as the learning of concepts) becomes the problem of didactics and its assessment.

School knowledge as "the things of logic" is the related to medieval rules of God in which there is an assumption that there is a universe of well-ordered concepts waiting for the proper investigation and expression. History, physics, mathematics, literature, and even the arts, are taken as having logical disciplinary structures which function as foundations from which learning is to occur.

Concepts divide, order and represent phenomena which are to be instilled in the person who wishes to examine phenomena. If we consider the "problem-solving" spoken about earlier in the "constructivist" pedagogues, we find the formulated skills of problem-solving are analytically described and separate from the concepts being described.

Why is the making of school subjects as "things of logic" significant when considering the tasks of a national evaluation? The pre-coding of the world as a purely logical "thing" in school curriculum is important when we think about the complexities of social life. When the "messiness " of our knowledge is obscured, so are the complexities and contingencies of knowledge that gives organization to our social, political and personal lives. The assumptions of school knowledge shape and fashion how questions are asked about social betterment and personal competence in the world. The social/ethical and political implications of the organization of knowledge are lost (see, e.g. Popkewitz, 1984).

The significance of the alchemy is also in the principles that order how teachers and students are to "learn" about who they are and what they are in their worlds-- I called this learning earlier as the forming of an identify and personal competence. Not only are grades achieved and the certification given. With learning concepts and information about science, social studies and mathematics are problem-solving methods to inquire, organize, and understand what the world and "self" are like. What is taken as knowledge and how to effect that knowledge (didactics) are technologies that direct how teachers and students are to reason about the at-large world and one's self in that world.

In this sense, we can view the imaginary subjects as positing a moral order through the rules of order, relations and identity. The moral order is both prior to and condition for the transmission of competence in the practices of teaching. Curriculum selection inscribes certain truths that are secure and enhance the well-being of social life through what are accepted as problems, questions, and responses in classroom practices. The moral order occurs as categories, distinctions, and differentiation of school subjects define purpose and direction through the classificatory systems applied. *The alchemy of school subjects construct a moral order which includes the child by excluding any social mooring for the child. The world of the child is seen as guided only by the internal logic of learning, individual motivation and personal "learning styles" and expression of an ahistorical "voice".*

SLEEPERS AND EPISTEMIC DRIFT

There is one further point that needs attention in the construction of an evaluation. It is the problem of **sleepers** and **epistemic drift** that occurs with state interventions. The very State activity to construct expertise in the evaluation can produce what Aant Elzinga * has called "epistemic drift". This occurs through a shifting of the types of questions and styles of investigation as external priorities and disciplinary practices interact in the construction of what is legitimate research. Although Elzinga was focusing primarily on science, we can think of the forming of a national evaluation as introducing new sets of legitimate actors and privileging certain epistemes in the research arena.

In the case of a national evaluation, there is a further peculiarity that occurs as the evaluation is formed. That has to do with the mobilizing of new actors who are sanctioned within the educational field. The testing and examination expertise is an example. It is clear in the Norwegian evaluation and State grants to groups to develop test development, that such a grouping of expertise is being produced and given institutional legitimacy through state practices. We can think about the testing group in Norway, for example, as having been "**sleepers**" within the educational field.[22] By that I mean that while certain types of testing expertise may have existed prior to the evaluation, that technical activity was not sanctioned as an activity of the state, and therefore little was "heard" from them. At this moment, that expertise is being "awakened" and quickly mobilized (with help of state funding) in the field of educational research.

The awakening of "sleepers" has a potential to influence the purpose and functions of research in education. Such mobilization has potential to shift the power relations in the field as graduate students are trained, centers established, and university positions created to practice this expertise.

The wakening of sleepers is neither "good", or "bad", by itself. It needs to be considered as occurring with a field with power arrangements. My argument throughout this essay has been that power in contemporary societies is deployed through the authorized systems of knowledge rather than through brute force. What is constituted as the sciences of schooling helps to form the objects of inquiry and "tells" us what is important and reasonable. Thus, the wakening of sleepers as a problem of epistemic drift is more than asking about what data to collect for an evaluation. Nor is the question of epistemic movement a question only of that science and its knowledge. It is also tied to the politics of knowledge and power in society.

SOME CONCLUDING THOUGHTS:

We can think of our social science categories as intellectual distinctions invented in the late 19th century and our cosmological assumptions as from the 15th century. Wallerstein (1991), in a book entitled *Rethinking Social Science*, argues that the social sciences maintain 19th century view of social relations that inscribe a universal conception of development (p.36). The purpose of inquiry was a search for universal laws and solutions to the "social question", carried in concepts such as the state-system and class, as well as in the distinct ideologies of conservativism, liberalism, and Marxism from which the normality of change as development was accepted. Further, the social sciences separated the nomothetic from the idiographic (p.2). In the current conjuncture, Wallerstein argues that the notion of universal development and the dichotomies of history and generalizability are "misleading and deconstructive". He suggests that there is a breakdown of certitude and the abandonment of the search for universal solutions. As well, the social sciences are no longer looking at the physical sciences as a model (dis-enchantment with the world) but to a re-enchanchment with the human construction of knowledge. A consequence, he argues, is that the presuppositions of 19th century social sciences, such as those of development and the distinction between ideographic and nomothetic, need to be rethought.

These presuppositions, once considered liberating of the spirit, serve today as the central intellectual barrier to useful analysis of the social world (p.1)

In many senses, the previous discussion of transformations in the arts, epistemology of science, economy, and politics point to the need to re-conceptualize the epistemological rules in the methodologies of research and evaluation. My intent in the discussion is to position the construction of a national evaluation within these larger historical changes and struggles whose conditions and categories cannot be taken for granted. It is also to problematize the categories of past social configurations -- words, if we focus on the educational arena as professionalization, measurement, and the dualism that makes the problem of governing as that of relation between the center and local.

Further, I believe that social and educational sciences need to historicize social practices, the approach that I took in this paper to consider the problem of evaluation. That historicizing, however, is not only of the actors in the educational arena but of the categories, distinctions and differentiations through which educational phenomena are ordered. This focus on knowledge is the sociology of knowledge. It is to recognize that words are historically formed within systems of ideas that inscribe styles of reasoning, standards and conceptual distinctions in school practices and its subjects. Further, we can think of the categories and distinctions of schooling as relating to an amalgamation of technologies, institutions and knowledges which form the power relations inscribed in contemporary reforms. Research and evaluation practices are responsive and a part of changing fields of regulation in which we, as part research communities and state agencies, participate.

The examples in this essay were presented as a strategy for considering what is typically taken-for-granted in the studies and evaluation of schooling; that is, the very systems of ideas about pedagogy, childhood, and achievement that normalize and discipline how schooling is interpreted and acted upon *are to be made into the objects of inquiry*. Yet the making of knowledge as the object of our questions about power is difficult because our distinctions seem natural and part of the doxa of the world -- unquestioned and unquestionable. That is, it seems that we "naturally" think of participation or children's development as good, and that our systems of measurement tell us whether we are achieving our "outcomes"; it is so "natural" to think in these ways that we do not recognize that the systems of ideas inscribe a double world. The norms of success imply its opposite, failure; with competence is incompetence; with concepts of achievement is deficiency. The norms of the opposite stand silent within our "seeing", "noticing", "attending to", and acting on as the "facts" of schooling itself.

To make the categories of knowledge in evaluation as a problem of self-scrutiny is to return to a theme that is threaded through the prior themes. Social science and evaluation are discourses which are both systems of interpretation and systems of action. This relates the problem of evaluation to the disciplining and regulatory functions of research itself. While reforms and the evaluation are to open the spaces in which teachers make decisions, *that space is disciplined and governed not only in the way policy is formulated, but through a cognitive structuring through which the individual participates and acts*.

Foucault (1979) called this "problem" as that of a "governmentality," that is, the tactics through which society was regulating as the patterns of personal decision-making and "reasoning" to judges individual competence and achievements was interwoven with macro problems of governing the State. In the 19th century, Foucault argues, there occurred a new relation between state governing practices and individual behaviors and dispositions.[23] Whereas medieval notions of "self" were determined through positions in a universal hierarchy with God at the apex, the modern person in liberal democratic societies had an individuality which likes to regard itself as autonomous and imagines itself as having sovereign rights determined within the limitations of geo-political boundaries. In one setting, reason was directed vertically towards divine domain of judgement rather than, in another setting, to an horizontal, bordered world with levels of autonomy and obligations. The shift in sovereignty in the 19th century included the forming of the modern, state responsible for the welfare of its citizens. In this context, the identity of individuals was linked to the administrative patterns found in the larger society. In multiple social arenas, intervention strategies of state reforms occurred as new institutions of health, labor, and education appeared alongside the emergence of the new social welfare goals of the state.[24] The mass schooling of the late 19th century and the discourses of school curriculum were linked to the administrative patterns of the state and a part of the construction of a "governmentality".

As I conclude, I hope that my argument is not construed as against the cohesion, standards and clarity of state goals, or as against commitments towards greater social equity through state inventions; neither am I the 19th century Luddite who fears all technologies and "things" of the modern and of the state as "evil". My point is very different. The changes of the present seem as "natural" and beneficial, the languages of educational reform are seductive; but in a world in which the changes occur are complex. That complexity cannot be assumed in evaluating school practices but must be part of the problem of investigation.

REFERENCES

Anderson, B. (1991). *Imagined communities: reflections on the origin and spread of nationalism* (Revised edition ed.). London: Verso.

Bachelard, G. (1984). *The New Scientific Spirit*. Boston: Beacon Press. (A. Goldhammer, Trans.)

Berger, P., Berger, B., & Kellner, H. (1973). *The Homeless Mind:: Modernization and Consciousness*. New York: Vintage.

Bernstein, R. (1992). *The New Constellation: The Ethical-Political Horizons of Modernity/Postmodernity*. Cambridge: MIT Press.

Bourdieu, P. (1984). *Distinction: A Social Critique of the Judgment of Tast*. Cambridge: Harvard University Press.

Bourdieu, P. (1990). *The Logic of Practice*. Stanford: Stanford University Press.

Bourdieu, P., & Passeron, J. (1977). *Reproduction in Education, Society and Culture*. Beverly Hills: Sage.

Boyer, R. (1989). The capital labor relations in OECD countries: From the Fordist "golden age" to contrasted national trajectories. In J. Schor, ed., Capital Labor Relations.

Bruckerhoff, C. (1991). *Between Classes: Faculty Life at Truman High*. New York: Teachers College Press.

Butler, J. (1992). Contingent Foundations: Feminism and the Question of "Postmodernism.". In J. Butler & J. Scott (Eds.), *Feminists Theorize the Political* (pp. 3-21). New York: Routlege.

Butler, J. (1993). *Bodies that matter: On the discourse limits of "sex"*. New York: Routledge.

Cangilheim, G. (1988). *Ideology and Rationality in the History of the Life Sciences*. Cambridge, MA: MIT Press. (Trans., Goldhammer, A.)

Carlgren, I. (In press). National curriculum as social compromise or discursive politics: Some reflections on a curriculum-making process. In Journal of Curriculum Studies.

Dahlberg, G., & Åsén, G. (1994). Evaluation and regulation: A question of empowerment. In P. Moss & A. Pence (Eds.), *A valuing quality*. London: Paul Chapman Publishing.

Donzelot, J. (1991). Pleasure in Work. In G. Burchell, C. Gordon, & P. Miller (Eds.), *The Foucault Effect. Studies in Governmentality* (pp. 251-280). Chicago: University of Chicago Press.

Dumm, T. (1993). The new enclosures: Racism in the normalized community. In R. Gooding-Williams (Ed.), *Reading Rodney King: Reading Urban Uprising* (pp. 178-195). New York: Routlege.

Elias, N. (1978). *The history of manners: The civilizing process* (Vol. 1). New York: Pantheon. (E. Jephcott, translator)

Elzinga, A. (1985). Research, bureaucracy and the draft of epistemic criteria. In B. Wittrock & A. Elzinga (Eds.), *The University Research System: The Public Policies of the Home of Scientists* (pp. 191-220). Stockholm, Sweden: Almquist & Wiskell International.

Englund, T. (1994a). Communities, markets and traditional values: Swedish schooling in the 1990s. *The Journal of Curriculum Studies, 2*, 5-29.

Englund, T. (1994b). Education as a citizenship right - a concept in transition: Sweden related to other Western democracies and political philosophy. *The Journal of Curriculum Studies, 26*, 383-399.

Fatis, S. (1992). Firms trim hierarchies, empower workers. *The Capital Times, Dec. 25, 1992*, 4B-5B.

Foucault, M. (1979). Governmentality. *Ideology and Consciousness, 6*, 5-22.

Giddens, A. (1987). *Social Theory and Modern Sociology*. Stanford: Stanford University Press.

Giddens, A. (1990). *The Consequences of Modernity*. Stanford, CA: Stanford University Press.

Goodwyn, L. (1978). *The Populist Movement: A Short History of the Agrarian Revolt in America*. Oxford, UK: Oxford University Press.

Granheim, M., Kogan, M., & Lundgren, U. (Eds.). (1990). *Evaluation as policymaking: Introducing evaluation into a National decentralised educational system*. London: Jessica Kingsley.

Hertfeld, M. 1992. *The Social Production of Indifference: Exploring the Symbolic Roots of Western Bureaucracy*. Chicago: The University of Chicago Press.

International Labour Organization. (1994). *Consequences of structural adjustment for employment, training, further training and retraining in the metal trades* (Report II ed.). Geneva: International Labour Office, Sectoral Activities Programme.

Interstate New Teacher Assessment and Support Consortium. (1992). Models standards for beginning teacher licensing and development: A resource for state dialogue; draft for comment/September 1, 1992. Washington, D.C.: Council of Chief State School Officers.

Jay, M. (1993). *Downcast eyes; The denigration of vision in twentieth-century French thought*. Berkeley: The University of California Press.

Johansson, I. (1993). Principles of legitimation in educational discourse in Iceland and the production of progress. *Journal of Educational Policy, 8*, 339-351.

Kallos, D., & Selander, S. (1993). Teacher education and teachers' work in Sweden: Reform strategies and professional reorientation. In T. Popkewitz (Ed.), *Changing Patterns of Power: Social Regulation and Teacher Education Reform* (pp. 211-262). Albany, New York: State University of New York Press.

Kantor, R. (1989). *When Giants Learn to Dance: Mastering the Challenge of Strategy, Management, and Careers in the 1990s*. New York: Simon and Schuster.

Ladson-Billings, G. (1994). *The Dreamkeepers: Successful teachers of African American children*. San Francisco: Jossey-Bass.

Ladson-Billings, G., & Tate, B. (1994). Toward a critical race theory of education. Madison, WI: The University of Wisconsin, Curriculum and Instruction, mimeo.

BEST COPY AVAILABLE

Lather, P. (1991). *Getting Smart: Feminist Research and Pedagogy Within the Postmodern*. New York: Routlege.

Latour, D. (1994). On technical mediation - Philosophy, sociology, genealogy. *Common Knowledge, 4*, 29-64.

Laursen, P. (1994). Teacher thinking and didactics: Prescriptive, rationalistic and reflective approaches. In I. Carlgren, G. Handal, & S. Vaage (Eds.), *Teachers' minds and actions: research on teachers' thinking and practice* (pp. 125-137). London: Falmer Press.

Lawn, M. (1994). The end of the modern in teaching? Implications for professionalism and work. In S. Lindbland & D. Kallos (Eds.), *New policy contexts for education: Sweden and United Kingdom* (pp. 169-190). Umea: Pedagogigiska institutionen, Umea universitet.

Lawn, M. (1995). *Restructuring teaching in the UK and USA: Towards the differentiated, flexible teacher*. Birmingham, UK: Westhill College. (Mimeo)

Lindblad, S., & Wallin, E. (1993). On transition of power, democracy and education in Sweden. *The Journal of Curriculum Studies, 25*, 77-88.

Lloyd, C. (1991). The methodologies of social history: A critical survey and defense of structurism. *History and Theory: Studies in the Philosophy of History, 30*, 180-219.

Lloyd, G. (1984). *The Man of Reason: "Male" and "Female" in Western Philosophy*. Minneapolis: University of Minnesota Press.

Lundahl, L. (1990). New variations on old themes: The Swedish conservative party and the battle over comprehensive education, 1900-1985. *Journal of Educational Policy, 5*, 157-166.

Manicas, P. (1987). *A History and Philosophy of the Social Sciences*. Oxford: Basil Blackwell.

McLaren, P. (1986). *Schooling as a Ritual Performance*. Boston: Routlege and Kegan Paul.

Merelman, R. (1976). On interventionalist behavioralism: An essay in the sociology of knowledge. *Politics and Society, 6*, 57-78.

Meyerson, A. (1992). IBM to eliminate 25,000 jobs in 1993 and shut plants. The New York Times, Dec. 16, 1992, P. 1.

National Council of Teachers of Mathematics. (1989). *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Novoa, A. (1993). The Portuguese State and Teacher Education Reform: A socio-historical perspective to changing patterns of control. In T.

Popkewitz (Ed.), *Changing Patterns of Power: Social Regulation and Teacher Education Reform* (pp. 53-86). Albany: State University of New York Press.

Ömholt, T. (1994). The regional and local perspective. Paper given at the Conference on the Development of Evaluation in a Decentralized Educational System. May 30-June 1, 1994, Lofoten, Norway.

Pateman, C. (1988). *The Sexual Contract*. Stanford, CA: Stanford University Press.

Popkewitz, T. (1984). *Paradigm and Ideology in Educational Research, Social Functions of the Intellectual*. London & New York: Falmer Press.

Popkewitz, T. (1990). Some problems and problematics in the production of evaluation. In M. Granheim, M. Kogan, & U. Lundgren (Eds.), *Evaluation as policymaking: Introducing evaluation into a National decentralised system* (pp. 103-118). London: Jessica Kingsley.

Popkewitz, T. (1991). *A Political Sociology of Educational Reform: Power Knowledge in Teaching, Teacher Education, and Research*. New York: Teachers College Press.

Popkewitz, T. (Ed.). (1993a). *Changing Patterns of Power: Social Regulation and Teacher Education Reform*. Albany, New York: State University of New York Press.

Popkewitz, T. (1993b). Professionalization in teaching and teacher education: Some notes on its history, ideology and potential. *Teaching and Teacher Education, 10*, 1-14.

Popkewitz, T. (1993c). U. S. teacher education reforms: Regulatory practices of the state, university, and research. In T. Popkewitz (Ed.), *Changing Patterns of Power: Social Regulation and Teacher Education Reform* (pp. 263-302). Albany, NY: State University of New York Press.

Popkewitz, T. (1993). Professionalization in Teaching and teacher education: Some notes on its history, ideology, and potential. In *Teaching and Teacher Education* (Vol. 10 (1), pp. 1-14). Great Britain: Pergamon Press Ltd.

Popkewitz, T., & Brennan, M. (In Press). *The Disciplining of teaching and the teacher: The contributions of Foucault to critical studies of schooling*. New York: Teachers College Press.

Popkewitz, T., & Lind, K. (1989). Teacher incentives as reform: Implications for teachers' work and the changing control mechanism in education. *Teachers College Record, 90*, 575-594.

Popkewitz, T., & Pitman, A. (1986). The idea of progress and the legitimation of state agendas: American proposals for school reform. *Curriculum and Teaching, 1*, 11-24.

Popkewitz, T., Tabachnick, B., & Wehlage, G. (1982). *The Myth of Educational Reform: A Study of School Responses to a Program of Change*. Madison: University of Wisconsin Press.

Probyn, E. (1990). Travels in the postmodern: Making sense of the local. In L. Nicholson (Ed.), *Feminism Postmodernism* (pp. 176-189). New York: Routlege.

Resnick, L., Levine, J., & Teasley, S. (1991). *Perspectives on Social Shared Cognition*. Washington, D.C.: The American Psychological Association.

Riley, D. (1988). *Am I That Name? Feminism and the Category of "Women" in History*. Minneapolis: University of Minnesota Press.

Rooney, E. (1989). *Seductive reasoning: Pluralism as the problematic of contemporary literary theory*. Ithaca: Cornell University Press.

Ross, D. (1991). *The Origins of American Social Science*. New York: Cambridge University Press.

Scott, J. (1991). The evidence of experience. *Critical Inquiry, 17*, 773-797.

Shon, D. (1987). *Educating the reflective practioner*. San Francisco: Josey-Bass.

Simola, H. (1993). Eduational Science, The State, and Teachers: Forming the corporate regulation of Teacher Education in Finland. In T. Popkewitz (Ed.), *Changing Patterns of Power: Social Regulation and Teacher Education Reform* (pp. 161-210). Albany: State University of New York Press.

Smith, M., & O'Day, J. (1990). Systemic School Reform. *Politics of Education Association Yearbook*, 233-267.

Tabachnick, R., & Zeichner, K. (In Press). Using action research to support conceptual change in the teaching of science. In M. Watt (Ed.), *Action research and the mathematics and science education: Purposes, practices and potential*. New York: Teachers College Press.

Tiles, M. (1984). *Bachelard: Science and Objectivity*. Cambridge, UK: Cambridge University Press.

Toews, J. (1987). Intellectual History after the linguistic turn: The autonomy of meaning and the irreducibility of experience. *The American Historical Review, 92*, 879-907.

Toulmin, S. (1990). *Cosmopolis, the Hidden Agenda of Modernity*. New York: The Free Press.

Wagner, P., Weiss, C., Wittrock, B., & Wollman, H. (Eds.). (1991). *Social sciences and modern states: National experiences and theoretical crossroads*. New York: Cambridge University Press.

Walch, J. (No date). *The commification of a social good: The case of Swedish educare*. Stockholm: Stockholm Institute of Education.

Wallerstein, I. (1991). *Unthinking Social Science: The Limits of Nineteenth-Century Paradigms*. Cambridge, UK: Polity Press.

Weiler, H. N. (1990). Decentralisation in educational governance: An exercise in contradiction? In M. Granheim, M. Kogan, & U. Lundgren (Eds.), *Evaulation as policymaking: Introducing evaluation into a national decentralised educational system* (pp. 42-65). London: Jessica Kingsley.

Weis, L., & Fine, M. (Eds.). (1993). *Beyond Silenced Voices: Class, Race, and Gender in United States Schools*. Albany: The State University of New York.

Young, R. (1990). *White mythologies: Writing, history and the West*. New York: Routledge.

Zerubavel, E. 1993. *The Fine Line: Making Distinctions in Everyday Life*. Chicago: The University of Chicago Press.

1. As I thought about this problems of this paper, I appreciated my discussions with a number of people; Gunilla Dahlberg, Lynn Fendler, Carl Jordell, Hannu Simola, Liz Lundahl. I also appreciated the comments from the seminar group in education at Umea University, the Department of Teacher Education at the University of Helsinki, and the participants in the seminar on educational research at the Swedish Skolverket.

2. My argument is not about which term to describe the changes occurring, but that such changes embody sets of circumstances that are sociologically different from previous efforts to administer social change and need different forms of representation to interpret than those that have been brought in from the past.

3. Someone might object to my use of democratic "instinct" that juxtaposes individual involvement and state monitoring. But I think that the modern notion of democratic is historically tied to a conception of the state that is to administer democratic processes.

4. There is also a challenge to the privileging of "vision" as the major sense organ for verifying knowledge. See, eg., Jay, 1994. While I focus on certain "tendencies" of knowledge, I recognize the debates and struggles about the interpretations and the normative commitments.

5. The difficulty in educational studies was that this issue was under-theorized and, overtime, "qualitative" studies tended to naturalize the subjective elements of schooling. There were, of course, counter-examples, such as Bruckerhoff, 1991; McLaren, 1984, and Weis & Fine, 1993.

6. see, e.g. Young, 1990; Toews, 1987; Lloyd, 1991). For general treatments of changes in the epistemological constructions of science and social science, see Toulmin, 1990; Cangilheim, 1988, Tiles, 1984; Bachelard, 1984; Manicas, 1987; for discussion about education, see Popkewitz, 1991; in one sense, the work of Resnick et al (1991) in psychology. Also see Elzinga's (1985) discussion of epistemic drift in the science, focusing on the relation of State and business policies to the forming and reforming of fields within science. The philosophical and sociological demise of Cartesian dualisms can be found in contemporary scholarship which focuses on the logic of science as not in the analytical formulations of information but in the ways that concepts and generalizations emerge from methods, and on the manner in which the questions and approaches of science are themselves socially produced within moral and political contexts.

7. The changes in the competencies of the teacher that I speak about here can also be understood as occurring in England. See, e.g., Lawn, 1994.

8. These changes in business organizations produce a revamping of such giant international corporations as General Motors, Sears, Roebuck, and I.B.M. The loss of 25,000 jobs within the giant computer company, IBM, reflects the changing world of work as smaller units with greater self-management are constructed (Meyerson, 1992). It is argued, for example, that the new model for computer and technological development involves alliances of small, more innovative companies, sometimes with government support and sometimes without. Every three years, the world micro-chip makers have been able to put four times as many transistors on a silicon chip, leading to vast increases in power and miniaturization. Each new generation of microchips created a new computer industry which overthrew the previous one.

9. The re-inscriptions of the individual in social spaces that is a central epistemological and political concern of post-modern literatures. Post-modern feminist literature, for example, seeks to de-stabilize domestic notions of differences through its explorations of the relation of biology and gender differences. In doing so, temporal and spatial dimensions of the concept of women are made problematic. This is accomplished by focusing on the manner in which gender is discursively and institutionally constructed and changes over time. The specific locales or contexts of women's lives are thus, re-positioned into those social spaces by which gendered relations are constructed (Probyn, 1990; Riley, 1988). The explorations are of how the "self" is differentially constructed through

discursive practices -- "how difference is established, how it operates, how and in what ways it constitutes subjects who see and act in the world" (Scott, 1991, p.777). (The significance of social space raises important questions in contemporary social and ethnographic research about how context is defined.)

10. Within new patterns of governing is the importation of two Anglo-American words--"curriculum" and "professionalism". The word, "curriculum", brings into focus distinctions about teaching, such as concerns with local planning, organizing, managing and evaluating of school knowledge. The calls for professionalism can be understood as also related to the creation a different occupational identity as teachers work with state mandates of goal directed pedagogy. Again drawing on particular Anglo-American concepts of professionalism, discourses of professionalism are received in Sweden to re-vision the work of the school to include greater teacher responsibility and flexibility in implementing goal governed approaches of the state (See, e.g. Popkewitz, 1993b).

It is interesting to note that many Scandinavian countries have a Germanic tradition in which the word, profession, tended not to be used in talking about educated occupations such as law or medicine. Also the strong state centered tradition tended to make the educated occupations tied more closely to the government with a less autonomous civil society. It is also important to note that hidden in discourses professions is its relation of the state, the development of capitalism, and issues of gender (See, e.g., Lather, 1991; Popkewitz, 1993a; Popkewitz and Lind, 1989).

11. In the United Kingdom, see Lawn, 1994.

12. Smith & O'Day (1990) suggest that we know what a successful school is ("research and common sense suggest that [the successful school] will have certain characteristics in common" (p.235). They will have

-a fairly stable staff
-enthusiastic and caring teachers who have mastery of subject matter and pedagogy;
-challenging curriculum that is integrated across grades and appropriate for diverse students;
-a depth of understand and complex thinking for children;
-A school climate conducive to learning that includes shared decision making and responsibility among staff and parental support and involvement (pp.235-236)

13. Populism has a long history in educational research and reform (See, Popkewitz, 1991).

14. The populism is one that appears in the late nineteenth century and a central element in the construction of educational discourses (see, e.g., Popkewitz et al, 1982; Popkewitz and Pitman, 1986).

15. In many ways, the more recent notion of the corporatist state embodies the idea of a "higher reason" as business and labor join through state action to produce social progress. We can also see a legacy of the distinction between state and civil society in current priorities of international funding agents. Recent guidelines of the World Bank and the International Monetary Fund juxtapose state decision-making with that of Non-Governmental Organization (NGO's) to foster the development of a civil society that will counter strong, centralist state traditions.

16. While stating this, actual practices to realize these commitments do vary in setting to reflect historical conditions through which political expression is given. The history of participation in Norway is different from, for example, that of the US.

17. I want to thank Lisa Visle of the University of Norway for giving me this distinction.

18. The historical work on the Swedish Conservative Party by Lundahl (1990) provides one approach which focuses on the relation of party to State that potential can be useful in exploring the changing patterns of the State, party, and governing that is forming.

BEST COPY AVAILABLE

19. These is an interesting parallel here with the Hegelian dialectic about the state/civil society distinction. Each assumes that a sameness can be achieved. Hegel's concern was with absolute progress and redemption; Marx inverted the idealism to materialism but maintained the same epistemological commitments that transformed the other into the negative of what is normalized as valued, thus making the other invisible. But this is another paper (See, Young, 1990).

20. For a non-psychological approach to problem-solving, see, e.g., Tabachnick & Zeichner (in press).

21. I appreciated this observation from a conservation with G. Ladson-Billings about current epistemological priorities in the study of the didactics of mathematics.

22. This metaphor was provided by Ulf P. Lundgren is a conversation that we had a few years ago about the "growth" of science.

23. The word "police" (and I assume later,"policy") was used to ensure a downward continuity between the ruler of the state and its populations. In the Middle Ages, governing was an extrinsic activity: the power of The Prince was to protect his geographic principality, with the question of regulation of souls left to the Church in preparing for an afterlife. By the 19th century, the meaning of governing involved a state as regulating and coordinating practices of individual behaviors and dispositions. The ensemble of institutions, procedures, analyses, reflections, calculations and tactics that define people as populations becomes paramount to the art of governing. The idea of social contract is made into a way of defining the mutual pledge of rulers and subjects.

24. These appearances of institution are an historical conjuncture rather than emerging from a grand design. It is coming together of the multiple developments in multiple arenas of social life that forms what I will later call a break or rupture.

Evaluation of a Prototype Teacher Enhancement Program on Science Performance Assessment[1]

Maria Araceli Ruiz-Primo, Richard J. Shavelson, and Gail. P. Baxter

Evaluation of a Prototype Teacher Enhancement Program to Transfer Performance Assessment Technology

Current science education reform addresses fundamental questions (e.g., Hurd, 1986) such as: What teaching methods enable students to understand the nature and culture of science? How can educators foster scientific literacy in students? How can science be related to everyday decision making? How can science understanding be assessed? The reform's answers to these questions are: Science instruction should parallel the methods used by scientists to understand the natural world (e.g., Raizen, Baron, Champagne, Haertel, Mullis, & Oakes, 1989). From this perspective, students have to do science--observe, hypothesize, record data, draw inferences and make generalizations--to solve scientific problems. By "doing" science students construct meaning both individually and in groups. Finally, assessment of student learning should parallel instructional reform.

Unless current assessment practice is changed, however, assessment will not parallel instruction reform, and reform in science education will not be comprehensively implemented in the classroom (e.g., Kulm & Stuessy, 1991; Shavelson, Carey, & Webb, 1990). Consequently, many states have responded with new policies which move achievement testing away from multiple-choice tests of basic skills toward performance-based assessments of knowledge and problem solving.

Changes in the nature and purpose of science instruction and subsequent changes in the nature of assessment exert pressure on the classroom teacher to change instructional and assessment practices (e.g., Shavelson & Baxter, 1990). Teachers are expected to shift from textbook and rote memory to constructivist teaching--teaching based on students' active construction of knowledge in problem-solving situations. To teach this way, teachers need to be well grounded in science to support an inquiry approach. They also need to change their role in the classroom from conveyors of facts and concepts to facilitators of knowledge construction. Moreover, they need skills in managing the physical and social organization of the classroom to support inquiry teaching (e.g., small groups of students working together). Finally, teachers need to have knowledge about new assessment policies and practices.

---

[1] Paper Presented at the AERA Annual Meeting, New Orleans, LA,April, 1994

To support teachers in the transition from traditional textbook teaching to constructivist teaching, a sustained program of in-service education is needed. Such a program would give teachers an opportunity to deliberate about the new perspectives in curriculum, teaching, learning, and assessment (e.g., Hurd, 1986; Shavelson, Copeland, Baxter, Decker, & Ruiz-Primo, 1994). In response to this need, the National Science Foundation (NSF) initiated in 1984 a Teacher Enhancement Program (TEP) to provide effective in-service education and foster the development and dissemination of improved models for conducting in-service education programs for science and mathematics teachers across the country (e.g., Fitzsimmons, Carlson, Burnham, Heinig, & Stoner, 1991).

One of the main goals of agencies like NSF is to have prototype in-service programs with significant impact not only on the sites where the program was developed but also in other settings (e.g., Fitzsimmons et al., 1991). Consequently, information about how these prototype programs were developed, what their characteristics are, and how they can be transferred to other sites has been requested not only by agencies like NSF but also by Congress (e.g., Fitzsimmons et al., 1991; Knapp, Shield, St. John, Zucker, & Sterns, 1988).

The purpose of this study was to conduct a formative evaluation of a prototype TEP which aims to provide teachers with the knowledge and skills needed to understand, use, and select science performance assessments. The formative evaluation tested the prototype program at different sites with a variety of facilitators and a variety of participants. It provided information about program components that need adjustment and revision to increase the probability of the program succeeding at different sites. Here we report on the transfer of a prototype TEP from the development site to outside sites.

Formative Program Evaluation

The practice of program evaluation has, once again, caught the attention of policy makers. They recognize that there is limited knowledge about the design and development of successful TEPs necessary for the reform of science education. What is needed is an understanding of the process involved in the development and implementation of successful in-service programs. Simply put, evaluative information is vitally needed.

Formative evaluation helps program developers better understand how and why the program is a success or a failure, to specify what aspects of the program are relatively more successful than others, and among which groups of participants (e.g., Cronbach et al., 1980). The main goal of formative evaluation is to modify and improve the design of any program while the program is still under development (e.g., Scriven, 1967) and therefore capable of being revised. Formative evaluation provides data as a basis for improving in-service programs (e.g., Chinien & Hlynka, 1993), helps to make judgments about how successful the program is (e.g., Guskey & Sparks, 1991), and helps to accumulate knowledge about how effective in-service programs are developed and adapted.

In spite of a general consensus among the policy makers and practitioners regarding the importance of formative evaluation, the great majority of in-service programs are still being implemented without prior formal evaluation (e.g., Knapp et al., 1988). If any kind of evaluation is carried out, it usually focuses either on changes in tests scores, or on information that is likely to be of tangential interest and utility to knowing and learning more about effective programs (See Ellis & Goulding, 1991; Ellis & Kuerbis, 1991; Gayford, 1987 as notably exceptions). The quality of in-service programs needs to be investigated, from planning and designing of the programs to follow-up of their impact.

The information provided from systematic formative evaluation may be used by developers and administrators for program improvement. Funding agencies and policy makers (e.g., NSF, Congress) may use the information to assist program developers with ways to formatively evaluate their program. For example, they can use the information to consider the implications of the evaluative information from any one in-service program for others also being operated by the same agency; or to disseminate how successful programs were developed, which may help other principal investigators who are leading the development or implementation of in-service programs.

An Approach to Formative Evaluation of TEPs

The approach proposed by Ruiz-Primo (1994) was used to carry out the evaluation of the prototype TEP. This approach makes two important assumptions. The first assumption is that a central task of evaluation is to facilitate the transfer of knowledge from some programs or sites to other programs or sites by explaining the processes that lead to the outcomes achieved (e.g., Cronbach, 1982.) The second assumption is that formative evaluation seeks to provide information to improve program performance by influencing immediate decisions about the program, especially about how its component parts and processes could be improved (e.g., Scriven, 1967, 1991a, b; Shadish, Cook, & Leviton, 1991).

This approach captures two types of information: (a) information related to the intrinsic value of the in-service program, and (b) information related to its potential dissemination. Both types of information help decision makers adjust and improve the program; however, each provides information on different aspects. Information about the intrinsic value refers to whether or not the program components (e.g., context, materials, and delivery conditions) are likely to meet the program goals. Information on the potential dissemination refers to how generalizable the program is to other settings (e.g., Weiss, 1972).

This approach to formative evaluation is built on three major elements: (a) the characteristics of the TEP to be evaluated; (b) the process involved in conducting the formative evaluation; and (c) the role and knowledge of the evaluator carrying out the formative evaluation (see Table 1).

TEP Characteristics. A TEP can be characterized as a system of interrelated *components*--context, goals, materials, delivery, and outcomes--which develop through three *stages of maturity*: (1) the *"planned program"*--the turn of an idea into a program for action; (2) the *"experimental program"*--a trial program to see what the program can accomplish, and (3) the *"prototype program"*--a model program that attempts to preview what will happen when the program is fully operational.

Evaluation Process. The formative evaluation process is conceptualized as an *iterative process* in which the program's goals are realized through successive approximations. The characteristics of the iterative process vary with the TEP's stages of maturity: from program reviews and revisions at the planned-program stage to program tryouts at different sites at the prototype program stage.

Congruent with the stages of maturity and the variations across successive trials (iterative process), the approach proposes variations in the evaluation process across the three *stages of formative evaluation*: (1) *"in-house reviews"* in which the evaluation provides information on the accuracy and adequacy of the planned-program materials, and on how feasible its operation is; (2) *"in-house tryouts"* in which the evaluation provides information on how the experimental program operates with typical facilitators and participants and what factors are associated with the program's success; and (3) *"outside tryouts"* in which the evaluation tests the prototype program in circumstances and with the population

TABLE 1
Characterization of an Approach to Formative Evaluation of TEPs.

TEP PROGRAM		EVALUATION PROCESS			EVALUATOR	
Stage of Maturity of the Program	Program's Components	Iterative Process: Conditions	Stage of Formative Evaluation:	Diversity of Methods of Evaluation		
Planned Program	Context Goals Materials Conditions of Delivery	Reviews-Revisions of the program before it is tried out.	In-house Reviews: Information about the accuracy and feasibility of the program.	Informal evidence (e.g., comments on content and possible difficulties).	Adaptable Role	Subject-Matter Expert
Experimental Program	Context Goals Materials Delivery Outcomes	Tryouts of the program with in-house staff and typical consumers. Revisions and reviews are also appropriate.	In-house Tryouts: Information about program operation: Variations and characteristics associated with more or less successful components.	Quantitative (e.g., small studies using quasi experimental or randomized designs) and qualitative methods (e.g., case studies)		
Prototype Program	Context Goals Materials Delivery Outcomes	Tryouts of the program in different sites with similar conditions to those proposed for operation. Revisions. Reviews are less necessary but may be appropriate too.	Outside Tryouts: Information on differences in delivery and effects from site to site, possible problems, and costs of implementation.	Quantitative and qualitative methods using research designs for estimating effectiveness are highly recommended		

that matched intended use when the program is fully operational. The approach stresses the use of different methods (quantitative and qualitative) and sources of information in the formative evaluation. These methods and sources vary as to the stage of formative evaluation, the information needed, and the audience requesting the information.

Evaluator's Role. Finally, for the formative evaluation to achieve its goal of improving a TEP, the approach assumes that the evaluator (1) possesses an *extensive knowledge* of the content of the program evaluated, and (2) is able to *adapt his her role* during the evaluation.

Comment. The approach recognizes that the development stages are not necessarily linear. There is always the possibility that at the experimental or prototype stage some components may have to return to a previous stage. In this way different components of the program may be at different stages of maturity at the same time (e.g., Cronbach et al., 1980). For example, during the evaluation of a prototype program, some activities may prove to be effective under all conditions, others may need minor adjustments while still others may have to be eliminated and new activities included to achieve the TEP's goals.

The Program Evaluated: A TEP to Inform about Science Performance Assessments to Teachers and Other Educators

As a part of a project funded by NSF (Shavelson & Baxter, 1990), a team at the University of California, Santa Barbara and the University of Michigan is in the process of developing two TEPs to transfer performance assessment technology to teachers and other educators. This section describes the characteristics of the TEP that were the focus of this formative evaluation.

The TEP is part of a larger project (Shavelson & Baxter, 1990) devoted to: (a) capturing the new technology involved in developing science performance assessments; (b) providing teachers and other educators with the knowledge and skills needed to understand, select, and use performance assessments embedded within the curriculum; and (c) training teachers and other educators to create and evaluate performance assessments.

The project is organized in an overlapping sequence of three Phases: Performance Assessment Technology, Training Development, and Field Test (Figure 1). In Phase I, Performance Assessment Technology, the emerging technology of creating performance assessments, is studied. The goals of this phase are to produce and evaluate performance assessments to be used as part of teacher pre-service and in-service education, and to make explicit the new technology's concepts and procedures so they can be transferred to teachers and other educators.

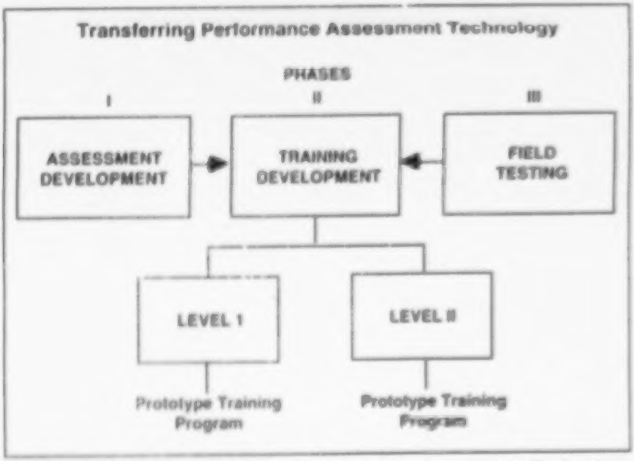

Transferring Performance Assessment Technology
PHASES
I
II
III
ASSESSMENT DEVELOPMENT
TRAINING DEVELOPMENT
FIELD TESTING
LEVEL 1
LEVEL II
Prototype Training Program
Prototype Training Program

Figure 1. A Project to Transfer Performance Assessment Technology to Teachers and Other Educators.

In Phase II, Training Development, a two level system of teacher enhancement is being developed. Level I training provides pre- and in-service teachers with knowledge and skills to understand, select, and use performance assessments. Level II training provides district personnel, teachers, scientist and other educators, working in teams, with the knowledge and skills needed to create and psychometrically evaluate alternative assessments. The major activities in Phase II are the development and evaluation of Level I and Level II training, culminating in prototype programs for field testing.

Phase III, Field Test, involves field testing both the Level I and Level II prototype programs to evaluate how well the training can be implemented in school districts with hands-on elementary science curricula, and the degree to which training meets its goals.

The project has two Principal Investigators with extensive experience in the development and evaluation of science performance assessments. For the development of Level I training the project has two instructional developers with extensive experience in developing teacher enhancement programs. The head of the development team is an expert in adult education. Level I training has also a

Coordinator who oversees all the activities related to the development of the program.

During the development of the TEP for Level I training a formative evaluation was carried out following the approach described above. This paper focuses on the third stage of development--outside tryouts with the prototype program over a nine-month period.

A program is becomes a prototype when the best possible program has been put together to meet its goals. The program can then be tried out in different sites, with different facilitators, and different participants. This stage may reveal a range of possible problems in operating the program on a large-scale (dissemination problems) and/or the components of the program that need to be modified in their delivery to improve effectiveness.

Characteristics of the TEP. The goals of the TEP are to provide pre- and in-service teachers with the knowledge and skills to: (1) understand the nature of assessment reform, (2) use these assessments in their classrooms, and (3) select existing assessments that are appropriate for evaluating individual student achievement or for monitoring the curriculum (Shavelson & Baxter, 1990). The mechanism for realizing these goals was a *prototype program* package that could be "exported" to school districts and be used by trainers (science educators) in those districts.

The TEP can be characterized as a *training approach* to staff development (Sparks & Loucks-Horsley, 1990). First, it is a workshop-type program in which the facilitator is the expert who establishes the content and flow of activities. Second, the training sessions are conducted with a clear set of objectives for learner outcomes. Third, the facilitator's role is to set the activities that will aid teachers in achieving the desired outcomes. This training approach is considered useful for realizing outcomes such as awareness, knowledge, and skills development, or when teachers require demonstrations of and practice on instructional techniques to be able to use the skills in their classrooms (e.g., Joyce, 1988; Joyce & Showers, 1980; Spark & Loucks-Horsley, 1990).

The Experimental Program. The experimental program stage of development was evaluated over five tryouts with project staff as facilitators and elementary science teachers, from the Science for Early Educational Development Project (SEED) in Pasadena, CA, as participants (see Ruiz-Primo, 1994; Ruiz-Primo, Baxter, & Shavelson, 1993). The evaluation collected information from a number of sources--documents, developers, participants, facilitators, evaluator--using a wide variety of methods--review of documents, direct observation, participants' products, responses to questionnaires, and interviews with facilitators.

The in-house evaluation revealed that participants acquired information they perceived as helpful to understand, use, and select performance assessments. The magnitude of the pre- and posttest program differences increased as the TEP's materials and delivery were improved based on the formative evaluation findings across the tryouts. The evaluation also revealed that facilitators' knowledge and experience in administering and scoring performance assessments were critical to the program achieving its goals (Ruiz-Primo, 1994; Ruiz-Primo, Baxter, and Shavelson, 1993).

These evaluation findings impacted the project in at least two ways. First, the original plan to "export" the prototype program as a package to other school districts was considered unrealistic. Second, it was clear that systematic training for facilitators was needed.

The Prototype TEP. After the five tryouts and many revisions of the program, the prototype TEP, the one evaluated here, had the following characteristics:

(a) The TEP's goals were three: understanding, use, and selection of performance assessments.

(b) The program reflected a hands-on instructional approach. Participants carried out hands-on elementary science performance assessments. With three of the assessments, they conducted the investigation and scored performance using procedure-based, evidence-based, and rubric (holistic) scoring systems. They conducted an exercise on interpretation of performance assessments scores, and another on selection of performance assessments. Figure 2 shows schematically the content for each goal.

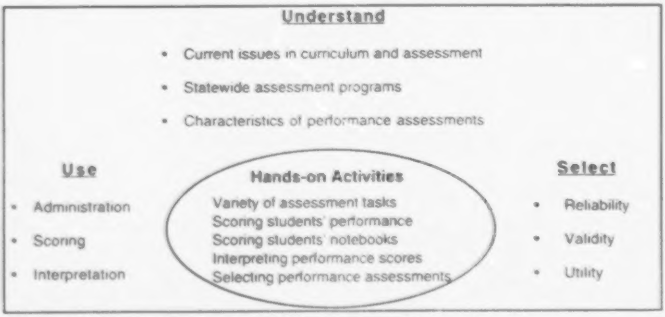

Understand
Current issues in curriculum and assessment
Statewide assessment programs
Characteristics of performance assessments
Use
Administration
Scoring
Interpretation
Hands-on Activities
Variety of assessment tasks
Scoring students' performance
Scoring students' notebooks
Interpreting performance scores
Selecting performance assessments
Select
Reliability
Validity
Utility

Figure 2. Workshop goals and the issues addressed.

(c) The TEP addressed 18 topics nested within one of the three goals: understanding (6 topics), use (5 topics), and selection (7 topics) of performance assessments. (Appendix A presents the topics by goal.)

(d) The program package included the "*Facilitators' Manual*", the "*Participant Notebook*", "*Nine Elementary Science Performance Assessments*", "*Transparencies*", and "*Videos*". The Facilitators' Manual is a detailed written script with the content, activities, and plan for delivering the program. For delivery purposes, the sequence and organization of content and activities were divided in "segments" (i.e., units, see Appendix B). The Participants' Notebook provides reduced copies of the transparencies used during the program and space for recording their notes and thoughts about each segment. Participants keep it for future reference.

(e) The program was designed to be delivered in 15 hours over three-days.

The Formative Evaluation Process

The stages of the formative evaluation of the Level I training program are presented in Figure 3. This study focused on the third stage: outside tryouts. The prototype TEP was tested in two sites with different facilitators and participants.

In this stage, the formative evaluation provided information on the adaptations needed to increase the probability of success when the program is fully operational. A central evaluation task, then, was to study how delivery and outcomes varied from site to site. Since the reproducibility of program results in different sites depends, in part, on how well the enactment of the TEP is described (e.g., Cronbach, 1982), evaluation findings also focused on identifying how the variations observed across sites were related to the characteristics of the program material and how these variations might be narrowed by adapting program materials.

With this perspective the evaluation of the prototype TEP focused on three components: *delivery, materials, and outcomes* (see Table 1). Program *delivery* refers to the conduct of the program with participants--how the content is conveyed to or constructed with the participants during the delivery. Program *materials* includes all documents that describe the program's content and activities, the sequence and organization of content and activities, and the delivery plan. Program *outcomes* refers to the participants' knowledge and skills about performance assessments acquired in the program.

Formative Evaluation Questions. The evaluation asked the following questions: (1) *Delivery*--Was the program delivered as it was designed so the program's goals can be achieved? (2) *Materials*--Which aspects of the materials led to major inaccuracies or variations during the delivery? and (3) *Outcomes*--Were the program's outcomes different from those found in previous tryouts?

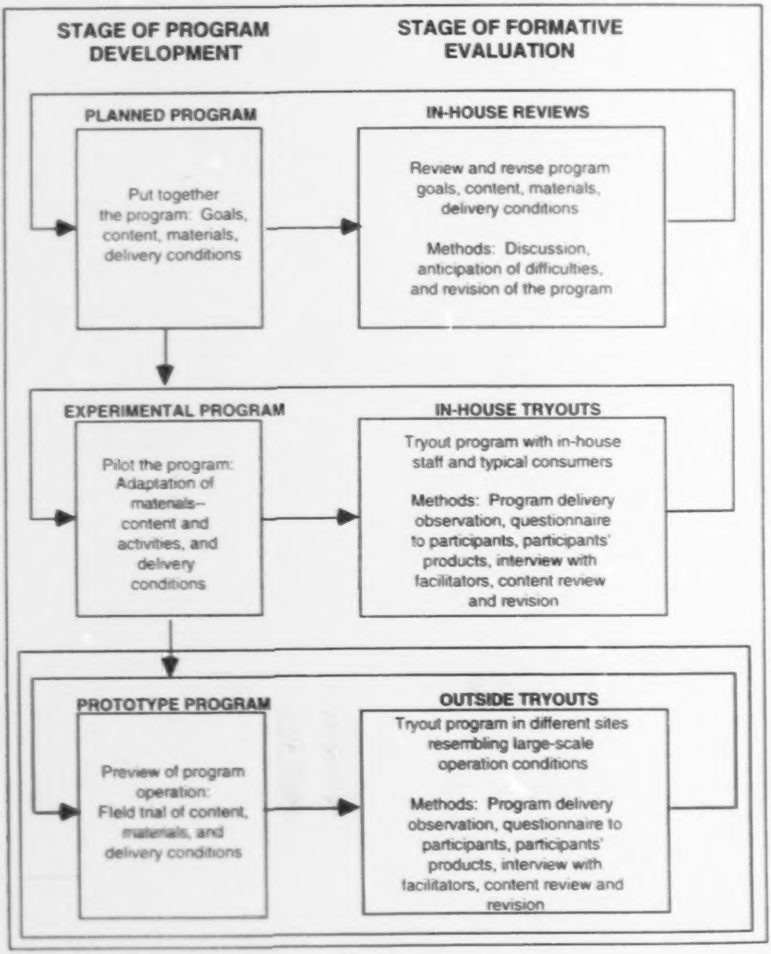

STAGE OF PROGRAM DEVELOPMENT
STAGE OF FORMATIVE EVALUATION
PLANNED PROGRAM
Put together the program: Goals, content, materials, delivery conditions
IN-HOUSE REVIEWS
Review and revise program goals, content, materials, delivery conditions
Methods: Discussion, anticipation of difficulties, and revision of the program
EXPERIMENTAL PROGRAM
Pilot the program: Adaptation of materials--content and activities, and delivery conditions
IN-HOUSE TRYOUTS
Tryout program with in-house staff and typical consumers
Methods: Program delivery observation, questionnaire to participants, participants' products, interview with facilitators, content review and revision
PROTOTYPE PROGRAM
Preview of program operation: Field trial of content, materials, and delivery conditions
OUTSIDE TRYOUTS
Tryout program in different sites resembling large-scale operation conditions
Methods: Program delivery observation, questionnaire to participants, participants' products, interview with facilitators, content review and revision

Figure 3. Strategy for formative evaluation.

. This evaluation, then, focused on whether the prototype program was implemented as expected in other sites with the same effects as those obtained where the program was developed.

The evaluation of program *delivery* centered on the characteristics of the "facilitators" and the enactment of the program's instructional methods. Information was collected on the: (a) facilitators' knowledge of program content--how accurately was the content delivered; and (b) implementation of the instructional plan--how adequate was the implementation.

Program *material* was evaluated as to how the characteristics of the content and the activities contributed to variations in the implementation of the program. Information was collected on the facilitators' perceptions of the program materials--content and activities and the instructional plan to deliver them.

The *outcomes* evaluation focused on program goals: to provide participants with the opportunity to "become familiar" with, not "experts" in, the nature, use and selection of performance assessments. Information was collected on the participants acquisition of knowledge about performance assessments.

Formative Evaluation Design and Instruments. The formative evaluation design followed from the Approach to the Formative Evaluation presented previously. Three outside tryouts were carried out at two sites each viewed as an iterative pilot study. This iterative process provides cumulative knowledge about the program which increases the program's robustness (e.g., Berk & Rossi, 1990).

For each tryout, the evaluation design called for collecting information before, during and after program delivery. The evaluation, then, took place before and after the delivery of the program, as well as in a pretest-posttest design during the delivery.

To provide a comprehensive view of the program as well as to cross check findings, different sources and methods of data collection were used. Evaluation data were collected from four sources: (1) documents, (2) facilitators, (3) participants, and (4) evaluator. Three data collection methods were used: (1) direct observation, (2) questionnaires, and (3) review of documents.

Table 2 presents a schematic representation of the formative evaluation design. This Table shows, for each component of the program evaluated, the sources of information, the instruments used to collect the data, and the point in time at which the instruments were administered during the tryout.

The evaluation of the program *delivery* used the evaluator as the main information source, and direct observation of the delivery as the main data collection method. To examine the program delivery, direct observation data were collected on each program topic. The delivery was videotaped and field notes were taken. The field notes included (a) time, (b) activity, (c) comments, and (d) suggestions. They were primarily descriptive, although the evaluators' reflections, interpretations, and direct suggestions made during the observation were also noted. Textual quotations were rarely included, paraphrasing was more typical. Participants served as a secondary information sources, responding to an "Opinion Questionnaire." This questionnaire used a Likert-type rating scale that elicited participants' perceptions about important topics and activities in the program. It also included open-ended questions that asked their opinion about the content and organization of the program.

TABLE 2

Design of the Formative Evaluation of the Program to Transfer Performance Assessment Technology

			Tryouts				
			Before	Delivery			After
Focus of the Evaluation	Source of Information	Method of Data Collection		Pre	During	Post	
Program Delivery	Evaluator	Direct Observation			X		
	Participants	Opinion Questionnaire				X	
Program Material	Documents	Review of documents	X				
	Participants	Opinion Questionnaire				X	
	Facilitators	Questionnaire					X
Program Outcomes	Participants	Self-report knowledge and skills questionnaire		X		X	
		Opinion Questionnaire				X	

Program *material* was evaluated using documents and facilitators as the main sources of data, and review of documents and questionnaires as the main data collection methods. Participants served as a secondary information sources to cross-check the findings of the other two sources. The evaluation of the material focused on the characteristics of the content and activities that allowed for variations on the implementation of the program by facilitators other than the project staff.

The Facilitators' Manual was considered the main document to be reviewed because it contained all aspects of the program: the sequence and organization of content and activities for the three days. These reviews also focused on the accuracy of the content and the adequacy of the instructional plan for delivering the content.

At the end of each tryout information on program material was collected from facilitators and participants. Facilitators responded to the "Facilitators' Critique Questionnaire" about the content and the activities of the program. The questionnaire used a Likert-type rating scale and open-ended questions that asked their opinion about the content and organization of the program material, and recommendations for changes in delivery, and the program as a whole. Participants' answers to the "Opinion Questionnaire" were used as a secondary source of information when necessary.

Program *outcomes* were evaluated using participants as the only source of information, and questionnaires as the method of data collection. The participants' knowledge of the content was evaluated by the "Self-Report Knowledge Inventory" in a pretest-posttest design. This inventory is a self-rating questionnaire that provides information about participants' knowledge of major topics covered in the workshop. Even though this type of instrument is not an achievement test, it has been shown to correlate highly with actual achievement, takes only a short time to administer, and is not threatening to teachers (see Tamir & Amir, 1981; Young & Tamir, 1977). The "Opinion Questionnaire" completed by the participants at the end of each workshop was also used as a secondary source of information.

Instruments were revised from one tryout to the next on the basis of their psychometric properties (when possible). Face validity was the criterion used to evaluate the validity of the participants' Self-Report Knowledge Inventory, and the Facilitators Critique and Opinion Questionnaires. Face validity--"the extent to which an instrument looks as if it measures what it is intended to measure" (Nunnally, 1970, p. 149)--is considered one of the best ways to facilitate decision makers' understanding of and belief in evaluation data (e.g., Patton, 1984). The instruments were revised by the developers and the principal investigators to increase their face validity (see Scriven, 1991a). Changes and adaptations were made to the instruments on the basis of developers' and principal investigators' comments, and the characteristics of the content. For example, new items were included in the Self-Report Knowledge Inventory when new topics were included in the content of the workshop.

Reliability was indexed by internal consistency. Reliability coefficients for the Self-Report Knowledge Inventory were obtained at both pretest and posttest on each tryout. Coefficients are presented when evaluation findings on the workshop outcomes are discussed.

Characteristics of the Outside Tryouts. Table 3 summarizes the general characteristics of the three outside tryouts. The characteristics include Facilitators' characteristics, participants, incentives for participation, duration, and data collection methods.

TABLE 3

General Characteristics of the Three Outside Tryouts of the Level I Training

ASPECTS	OUTSIDE TRYOUTS		
	SITE 1		SITE 2
	1	2	3
Facilitators' Characteristics	Both were Elementary Teachers and Staff Developers from the County School Office	Same Facilitators as Occasion 1	Two Resource Teachers from the District and One Elementary Teacher
Participants' Characteristics	Elementary Teachers	Elementary Teachers	Elementary, High School, and Resource Teachers
Participants' Incentives	Ps paid for the workshop and got one unit credit	Ps paid for the workshop and got one unit credit	None
Organization of the Program in Days	3 in a row	2 evenings and one full-day	2 full-days
Methods of Data Collection			
• Facilitator's Critique Questionnaire		√	√
• Direct Observation	√	√	√
• Review of Documents	√	√	√
• Self-Report Knowledge Inventory	√	√	√
• Opinion Questionnaire	√	√	√
• Descriptive Information Questionnaire	√	√	√

The program was piloted at two sites: Site 1--Southern California Superintendent of Schools, and Site 2--Middle Arizona School District. At Site 1 the program was piloted on two occasions with one elementary and one high school teacher as facilitators on both occasions. Participants were in-service teachers from different school districts in the county. At Site 2, the program was piloted once with two science resource teachers and one elementary teacher as Facilitators. Participants were in-service teachers and resource persons from the District.

At Site 1, participants paid a fee to the County Office to participate and get one unit credit course. At Site 2, participants were given time off from work. The first outside tryout, Site 1 occasion 1, was carried out in a three-day session. The second tryout was carried out in two evenings and one full-day. The third tryout at Site 2, was delivered in two full days.

The only data collection method not used on the first tryout was the Facilitator Critique Questionnaire. However, information was obtained from in-depth interviews carried out with both facilitators by one of the developers of the program. For the two remaining tryouts, all sources of information listed in Table 3 were used.

Facilitators Characteristics. Table 4 presents information on the characteristics of the facilitators. Facilitators from Site 1 were selected by the County Superintendent's Office. Both were females with 8.5 years experience, on average, as teachers and 3 as facilitators on different in-service programs. They felt they had adequate experience in hands-on science teaching and performance assessment, although information from direct observation during the delivery revealed that this might not be the case. Facilitator 1's background was education (i.e., elementary education and curriculum and instruction), whereas Facilitator 2 clearly had a science background (i.e., Zoology and Biological Science).

Both facilitators were informally "trained" on the two occasions by one of the program developers. One of the facilitators, F1, experienced the program as a participant in a previous implementation delivered by the project's staff. Facilitators received the Facilitators' Manual about two weeks before the tryout. Then, facilitators and the developer met to discuss the content and the logistics of the program for approximately 16 hours before the delivery. The developer walked them through the program (e.g., the sequence, the activities, implementation, use of transparencies) based on previous implementations. The developer also answered questions the facilitators had about the content. Finally, the developer met with the facilitators at the end of each session during the

TABLE 4
Characteristics of the Facilitators By Site

CHARACTERISTICS	OUTSIDE TRYOUTS				
	SITE 1		SITE 2		
	F[a]1	F2	F1	F2	F3
Years of Teaching Experience					
Elementary	7		28		6
Junior High		6		3	
High School		10		20	
Years of Science Hands-On Teaching					
Elementary	7		26		6
Junior High		6			
High School		10		20	
Years as Trainer of Science Teachers	4	2	9	6	3
Undergraduate Major					
Education	√				√
Science		√		√	
Other			√		
Advanced Degree					
MA Education	√		√	√	√
MA Science		√			
MA Other					
Ph. D. Education					√
Experience with Hands-On Science Teaching					
Novice					
User		√		√	NA[b]
Expert	√		√		
Experience with Science Performance Assessments					
Novice			√		
User				√	√
Expert					

[a] F = Facilitator.
[b] NA = No Answer

implementation. At these meetings, developer and facilitators discussed the content to be delivered during the next session.

Facilitators at Site 2 were males. They worked at a School District well recognized for its science curriculum and hands-on approach to teaching. Two of the facilitators, F1 and F2, have been resource teachers in the Districts' Science Resource Center for approximately 17 years. They had extensive experience as hands-on science teachers (23 year on average) and as trainers of teachers in the same district (9 years, on average). They also had experience in administering and scoring performance assessments. Facilitator 3 was chosen by the District for two reasons: They wanted him to be involved in the development of performance assessments for the Resource Center's curriculum units, and to be a trainer of teachers in the District on the use of performance assessments.

These facilitators were also informally "trained" by the Coordinator of the Level I training project. The three facilitators were participants in two previous implementations of the program. The first was delivered by the project's staff (see above) eight months before the tryout. The second was specifically arranged to prepare them to facilitate the training at their home site. They had the Facilitators' Manual and met with the Coordinator for approximately 6 hours to discuss program content and a plan to deliver it. During the implementation of the tryout, the Coordinator also met with the facilitators at the end of the first day to discuss concerns about the content to be delivered at the next session.

Participant Characteristics. Table 5 presents participants' characteristics across tryouts. Participants differed from site to site. Whereas Site 2 has been recognized as an exemplary hands-on school district, Site 1 is starting to move to a hands-on instructional approach. At Site 1, some teachers in some school districts were already using this approach, while others were not familiar at all with this new way to teach science.

TABLE 5

Characteristics of the Participants Across the Outside Tryouts

CHARACTERISTICS	OUTSIDE TRYOUTS		
	SITE 1		SITE 2
	1	2	3
Number of Participants	38	20	22
Mean Years of			
Elementary Teaching	13.29	10.93	6.09
Junior High Teaching	0.93	50	7.52
High School Teaching	0.38	96	9.56
Mean Years of Hands-on			
Elementary Teaching	6.9	7.81	5.57
Junior High Teaching	0.45	.06	5.38
High School Teaching	0.38	0.0	0.19
Undergraduate Major			
Education	9	2	9
Science	1	1	11
Other	21	13	2
Advanced Degree			
MA Education	7	4	10
MA Science	0	0	5
MA Other	4	1	1
Ph. D. Education	0	0	1

Most participants at both sites were elementary teachers. A few of them, particularly at Site 2, were junior high and high school teachers. Forty percent of the participants held a master's degree, 26 percent in education. Only one participant held a Ph.D. in education.

Evaluation Findings

This evaluation focused on three major questions about the prototype program: (1) "Is the program delivered as it was designed so the program's goals can be achieved?"; (2) "Which characteristics of the program's material lead to major variations across the sites?"; and (3) "Do the program's outcomes differ from those found in previous tryouts?"

Data were brought to bear on each program component--delivery, materials, and outcomes--for the three outside tryouts. To examine program delivery and materials, data were collected on each of the topics that constituted the program (see Appendix A). Data bearing on the outcomes component were based on pretest-posttest scores from the Self-Report Knowledge Inventory and the participants' opinions.

First, a summary of the evaluation findings across tryouts is presented along with the major decisions made about the content and delivery of the program. Next, two examples of the evaluation findings on program delivery and program material are presented.

Summary of Evaluation Findings

From revisions made to the experimental program during the five tryouts of the in-house evaluation, the TEP program was considered to be ready for implementation at other sites--the content and the instructional plan were adequate and the program had proven effective in achieving its goals. However, the evaluation findings that emerged through the outside tryouts revealed that the program still needed some adaptations to increase the likelihood of successful transfer to other facilitators at other sites. These findings are presented for each of the components evaluated.

Program Delivery. Evaluation during the delivery of the program produced four major findings: (a) The delivery of the program was modified by facilitators on all three tryouts. Modifications were: (1) "Superficial"--facilitators added new activities not directly related to the content or the goals of the program (e.g., to give prizes to participants, play games). (2) "Process"--facilitators modified the instructional plan in delivering some topics. These modifications ranged from minor modifications (e.g., reducing duration of an activity) to major modifications (e.g., changing the original instructional plan completely). And (3) "Content"--facilitators ignored or added topics/activities during delivery (e.g., omission of the research findings during the presentation of the technical qualities of performance assessments.)

The most obvious impact of superficial modifications was the change in the schedule for implementing the program. For example, in tryout 1, there was not enough time to discuss some topics or carry out some activities on the last day. Information collected from direct observation revealed that some of the process modifications led to discussions or raised questions that facilitators could not address properly. For example, on the first tryout at Site 1, facilitators were not able to close off discussions or to comment accurately on participants' accurate/inaccurate statements. At Site 2, process variations impacted program material. For example, by changing the instructional plan, facilitators had to change the "participants notebook". Finally, the evaluation found that the variations in the implementation of the program did not impact the achievement of the program's goals.

(b) Consistently across tryouts, the information delivered for certain topics was inaccurate (e.g., procedure-based scoring systems). Probably these topics were too complex for Facilitators to come to understand by reading the Facilitators' Manual. More emphasis should be paid to technical topics during facilitator training.

(c) Facilitators' knowledge and background were found to be a key to the success of program delivery. The Facilitators' knowledge about hands-on science instruction, their experience with the performance assessments used in the program (e.g., experience in administering and scoring performance assessments), their background in developing performance assessments, and their knowledge of the program content were all key factors in the quality of the delivery. For example, based on direct observation of delivery, we found that facilitators had difficulty identifying participants' misconceptions about performance assessments or answering "non-scripted questions" due to limited knowledge about and experience with hands-on instruction and performance assessments.

Direct observation also revealed that Facilitators tended to "know" only those parts of the manual they delivered; the rest of the content was not included in "their program." Facilitators at Site 1 displayed little understanding of the agenda and the content to be presented for each succeeding session. For example, on occasion 1 at Site 1, one participant asked facilitators if they would have the opportunity to watch a video in which performance assessments were administered to a whole class. Facilitators said "no" even though there was a video on this issue that participants would see the next day. Facilitators had not read the whole manual or previewed the videos included in the program package.

As might be expected, this situation influenced the quality of the delivery. For example, facilitators, instead of closely monitoring small group discussions/activities, tended to use that time to read the next topic in the manual. Yet, hearing participants' discussions and viewing their performances would help facilitators discover misconceptions or problems in understanding.

(d) Observation of the delivery revealed that the certain characteristics of the participants made a difference in the type of questions, discussions, and even level of enthusiasm about the topics presented in the program. Participants from Site 2, already aware of the need for alternative assessments in their hands-on science curriculum, were more focused in their discussions and the questions they posed to facilitators, and clearly more enthusiastic.

Findings from the program delivery evaluation called for some changes on the TEP, specially after tryout 1: (a) The content of the program needed to highlight the importance of certain critical topics to the facilitators. (b) Background material from popular professional journals for teachers and administrators needed to be provided (e.g., special issues of Educational Leadership). Articles from research journals did not motivate facilitators to read about performance assessments. And most importantly, (c) facilitators needed more thorough training on the content of the program before they could "own" it and deliver it successfully.

Program Materials. Major changes in the program materials were made after tryout 1. Three factors influenced these changes: facilitators' recommendations about the program, the variations observed during implementation, and an executive decision.

Facilitators made the following recommendations: the program needed more group discussion and more time for participants to "process" the information; the program included too much information, "less is more" according to one of the facilitators; the presentation of the technical qualities of performance assessments did not require as many charts (i.e., graphs) as those included; and the format of the manual needed improvement.

Information from the facilitator interviews conducted by one of the program developers revealed some reasons that facilitators had for modifications they made during the implementation. For example, they omitted some topics for two main reasons: (a) they considered the information irrelevant for teachers (e.g., research findings on the technical characteristics of performance assessments), or (b) time constraints. They included many new activities (e.g., "icebreakers", "sponge activities", "carousels", "prizes") because they wanted to give participants the opportunity to get up and move around and discuss ideas with each other, to keep them involved in the activities, and to provide positive feedback.

After the first tryout at Site 1, the principal investigator of the project made an "executive decision" to let program developers change the "format", not the "content" of the program. Because the content had already proven successful in achieving the program's goals, the developers' task was to enhance the "teacher-friendliness" of the program.

Discussions about changes in the instructional were the every day story in the project. Moreover, with changes in the plan came reviews and revisions of the content to insure that its accuracy was intact after modifications. Finally, a new version of the program material was tried out on the second occasion at Site 1, and at Site 2. Although the content remained almost the same, the instructional plan for delivery was substantially modified.

The new version of the program had the following characteristics: (a) Some topics were dropped from the program (e.g., two types of validity, curriculum sensitivity and discriminant validity), and others were reduced in scope (e.g., interrater reliability and intertask reliability). (See Appendix C for a list of the topics included in the program.) (b) New instructional activities were included to communicate content (e.g., mini-lectures; participant content checks.) (c) The instructional plan for delivering some topics was completely modified. For example, to motivate the review of the technical qualities of performance assessments prior to the selection exercise, the plan called for participants to improvise skits, in small groups, which depicted a specific technical quality (e.g., interrater reliability), act it out, and let the other participants guess the term depicted. (d) The content sequence was changed (see Appendix B.) (e) Characteristics of the Facilitators' Manual were changed (e.g., shaded boxes were used to indicate important points to make to the participants; mini-lectures were highlighted by a "box" and change in type font and size.) (f) The Participants' Notebook was modified based on changes in program content.

Reviews of the Facilitators' Manual during the modification period revealed new content inaccuracies as a result of the modifications (e.g., developers used inaccurate terms) and some instructional inadequacies (e.g., developers wanted participants to "fill-in the blanks" as a check on content knowledge, an activity that contradicted the hands-on philosophy of the program.)

Information from the Facilitators Questionnaire showed that, for facilitators, the TEP was effective in meeting its goals, and at an adequate level for the participants. In general, they thought that most of the segments were effective in meeting their particular objectives. However, all facilitators agreed that Segments 4, "Paper Towels Investigation", and Segment 5, "Introduction to Scoring Systems", were only "somewhat effective." Information collected from direct observation during the delivery corroborated this finding. Facilitators on all three tryouts had problems delivering these two segments.

Evaluation findings from the last two tryouts revealed that the "New Facilitator's Manual" still need improvement. Topics that proved to be consistently inaccurately delivered need to be revised to help facilitators deliver the content more accurately. Also, training need pay particular attention to these topics in the future.

Moreover, some of the "new activities" included in the modified manual should be presented as "optional" for facilitators. For example, according to facilitators from Site 2, some of the activities may not be appropriate for high school teachers (e.g., the improvisation [illegible]ds.)

Finally, for facilitators to acquire the "whole" picture of the program, they needed to spend a considerable amount of time and effort studying and learning the material. There is no doubt that the amount of information contained in the Facilitators' Manual is a contributing factor. Facilitators felt overloaded with the information, which may have influenced the effort they put into "owning" the program. However, the commitment that facilitators have to deliver a program with good quality is also an important factor. Facilitators at Site 2 were clearly more engaged with the delivery of the program. They met as a team at least four times before implementing the program to discuss content, modifications, and how the modifications impacted the delivery (e.g., schedule of the implementation, changes to the participants' notebook). These meetings lead facilitators to know better the characteristics of the program. In sum, delivery at Site 2 is a good example of the difference that facilitators' commitment can make.

Program Outcomes. The participants' self-reported knowledge about the topics addressed in the program and their opinions about the program were used as a source of information on outcomes.

The Self-Report Knowledge Inventory was administered to participants in a pretest-posttest design at each tryout. Designed to assess knowledge and skills acquired during the program, this instrument was somewhat unique. It had the appearance of a questionnaire and asked participants to indicate their understanding of the topics covered in the program.

To examine the differences in the knowledge and skills acquired by participants as a result of the program, a series of dependent t-tests for differences between pre- and posttest mean scores was carried out for each tryout. Table 6 presents the descriptive statistics and the reliability coefficients (i.e., internal consistency) at pre- and the posttest.

TABLE 6
Statistics for the Pre- and the Posttest Total Score

	1		2		3	
	Pre	Post	Pre	Post	Pre	Post
n participants	38		20		22	
Maximum	48		64		68	
Mean	19.20	38.22*	30.60	51.82*	31.57	56.47*
S. D.	6.40	5.15	6.26	7.08	8.72	7.52
Reliability	.91	.86	.87	.94	.95	.93

* Significant difference between Pre- and Posttest ($\alpha = .001$).

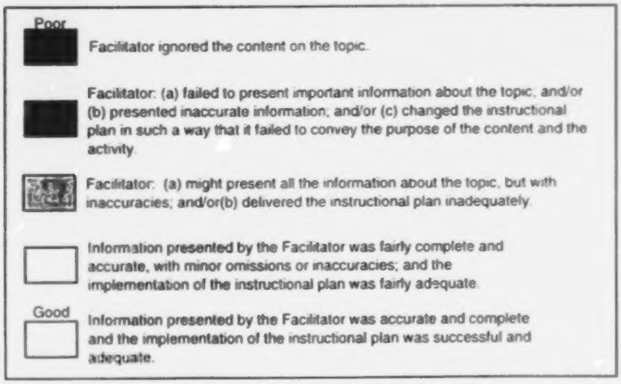
Significant differences between the pre- and the posttest mean total scores were observed across tryouts ($t(29) = 15.048$, $p < .01$; $t(16) = 9.78$, $p < .01$; $t(19) = 14.26$, $p < .01$, respectively.) Reliability coefficients for total scores were high and roughly of the same magnitude at both pre- and posttest across the three tryouts. Patterns of differences between the pre- and posttest were similar to those observed on the last two in-house tryouts.

Based on participants report of their knowledge, the program was effective in achieving its goals despite program modifications made by the facilitators during the delivery.

Examples of the Evaluation Findings

In this section we present two concrete examples of evaluation findings. One addresses program delivery findings and the other content findings.

Program Delivery. To examine program delivery, data were collected on each of the topics in the program (see Appendix A & C). Findings on the delivery of one topic across the three tryouts are presented concisely in a flowchart (Figure 4). The shaded boxes symbolize the quality of delivery--the accuracy of content and adequacy of the plan for delivering it. The darker the box, the poorer the delivery quality. A light box indicates that the desired quality was achieved. The criteria used to shade the boxes based on the quality of the delivery is presented in Figure 5. Thick-line boxes refer to the evaluation findings; square boxes refer to the evaluator's recommendations, and arrows represent the time sequence.



Poor

■ Facilitator ignored the content on the topic.

■ Facilitator: (a) failed to present important information about the topic; and/or (b) presented inaccurate information; and/or (c) changed the instructional plan in such a way that it failed to convey the purpose of the content and the activity.

▨ Facilitator: (a) might present all the information about the topic, but with inaccuracies; and/or(b) delivered the instructional plan inadequately.

☐ Information presented by the Facilitator was fairly complete and accurate, with minor omissions or inaccuracies; and the implementation of the instructional plan was fairly adequate.

Good

☐ Information presented by the Facilitator was accurate and complete and the implementation of the instructional plan was successful and adequate.

Figure 4. Criteria used to evaluate the quality of the delivery and shade the boxes in the flowcharts.

As an example, consider the delivery of the topic, "Intertask Reliability," of performance assessments. Figure 4 presents the evaluation findings from direct observation bearing on the delivery of this topic across three tryouts.

This topic dealt with technical characteristics of performance assessments (i.e., reliability, validity, utility). Its purpose was to make participants aware of the importance of considering the consistency of students' performance-assessments scores across different tasks.

The evaluation on the first and the last tryouts showed that the quality of delivery was poor. Although the inaccuracies in the information delivered to participants were different, both reflected the fact that facilitators needed more information and better understanding of this topic in order to deliver it accurately. During the second tryout, the quality of the delivery improved. Two factors influenced this improvement. First, Facilitator 2 delivered this topic on both occasions. Second, on occasion 2, the facilitator discussed the topic with the Coordinator of Level I training before it was delivered. Both factors, the experience in delivering the topic and the discussion, may have helped.

Participants' perceptions were used to triangulate on the findings from direct observation. Participants in tryout 1 were asked, "What recommendations would you make about the organization and the content of this workshop that you

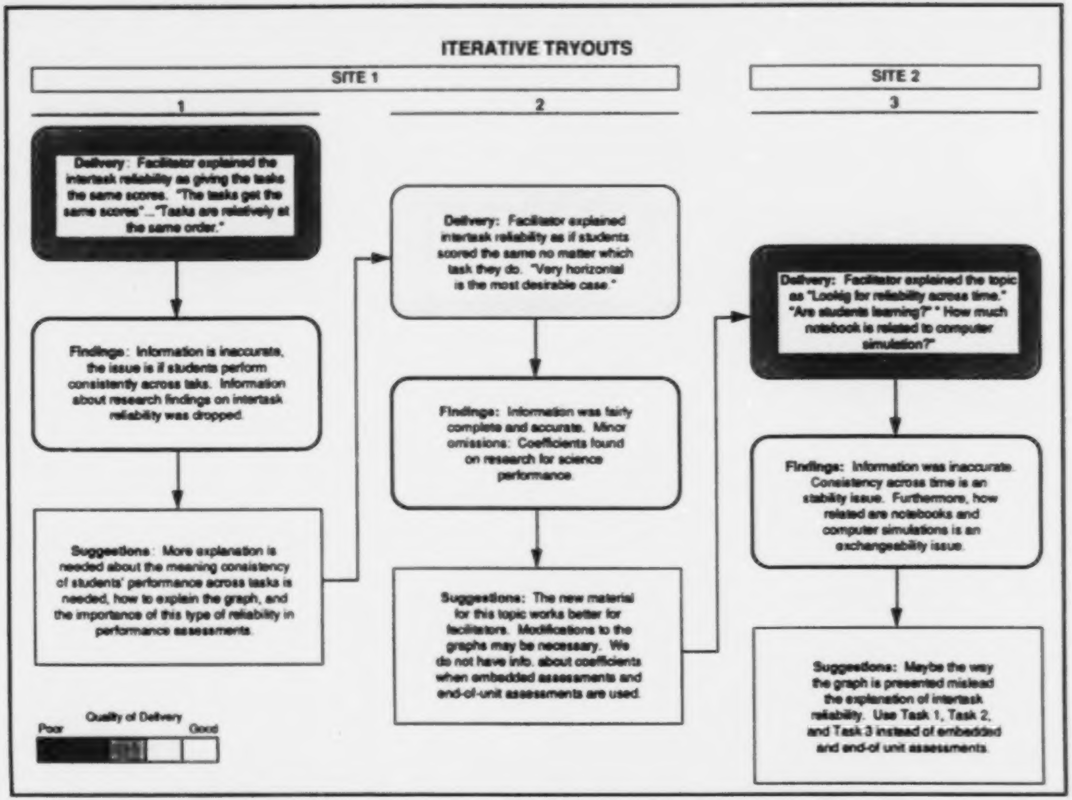
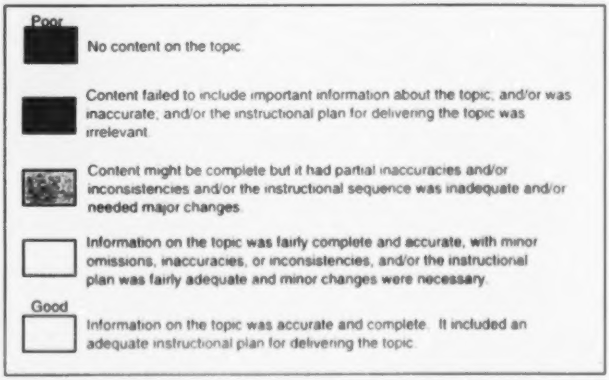

ITERATIVE TRYOUTS
SITE 1
SITE 2
1
2
3
Delivery: Facilitator explained the intertask reliability as giving the tasks the same scores. "The tasks get the same scores"..."Tasks are relatively at the same order."
Findings: Information is inaccurate, the issue is if students perform consistently across taks. Information about research findings on intertask reliability was dropped.
Suggestions: More explanation is needed about the meaning consistency of students' performance across tasks is needed, how to explain the graph, and the importance of this type of reliability in performance assessments.
Delivery: Facilitator explained intertask reliability as if students scored the same no matter which task they do. "Very horizontal is the most desirable case."
Findings: Information was fairly complete and accurate. Minor omissions: Coefficients found on research for science performance.
Suggestions: The new material for this topic works better for facilitators. Modifications to the graphs may be necessary. We do not have info. about coefficients when embedded assessments and end-of-unit assessments are used.
Delivery: Facilitator explained the topic as "Lookig for reliability across time." "Are students learning?" " How much notebook is related to computer simulation?"
Findings: Information was inaccurate. Consistency across time is an stability issue. Furthermore, how related are notebooks and computer simulations is an exchangeability issue.
Suggestions: Maybe the way the graph is presented mislead the explanation of intertask reliability. Use Task 1, Task 2, and Task 3 instead of embedded and end-of unit assessments.
Quality of Delivery
Poor
Good

Figure 5. Delivery of the topic "Intertask Reliability of Performance Assessments."

think would help to improve it?" Some participants recommended dropping the topics related to statistical issues (e.g., "get rid of statistical information on validity and reliability"; "leave out the statistical lesson"). Furthermore, two participants wrote that it was clear that facilitators did not know enough about the statistical information.

Future training of other facilitators, then, should provide a more detailed explanation of the meaning of consistency across tasks as well as the relevance of this topic in the context of performance assessment.

Program Content. Evaluation findings on program material were based on careful review and revision of the Facilitators' Manual, Participants' Notebooks, and Transparencies. The criteria used to shade the boxes based on the quality of the content is presented in Figure 6.



Poor

■ No content on the topic.

■ Content failed to include important information about the topic; and/or was inaccurate; and/or the instructional plan for delivering the topic was irrelevant.

▨ Content might be complete but it had partial inaccuracies and/or inconsistencies and/or the instructional sequence was inadequate and/or needed major changes.

☐ Information on the topic was fairly complete and accurate, with minor omissions, inaccuracies, or inconsistencies, and/or the instructional plan was fairly adequate and minor changes were necessary.

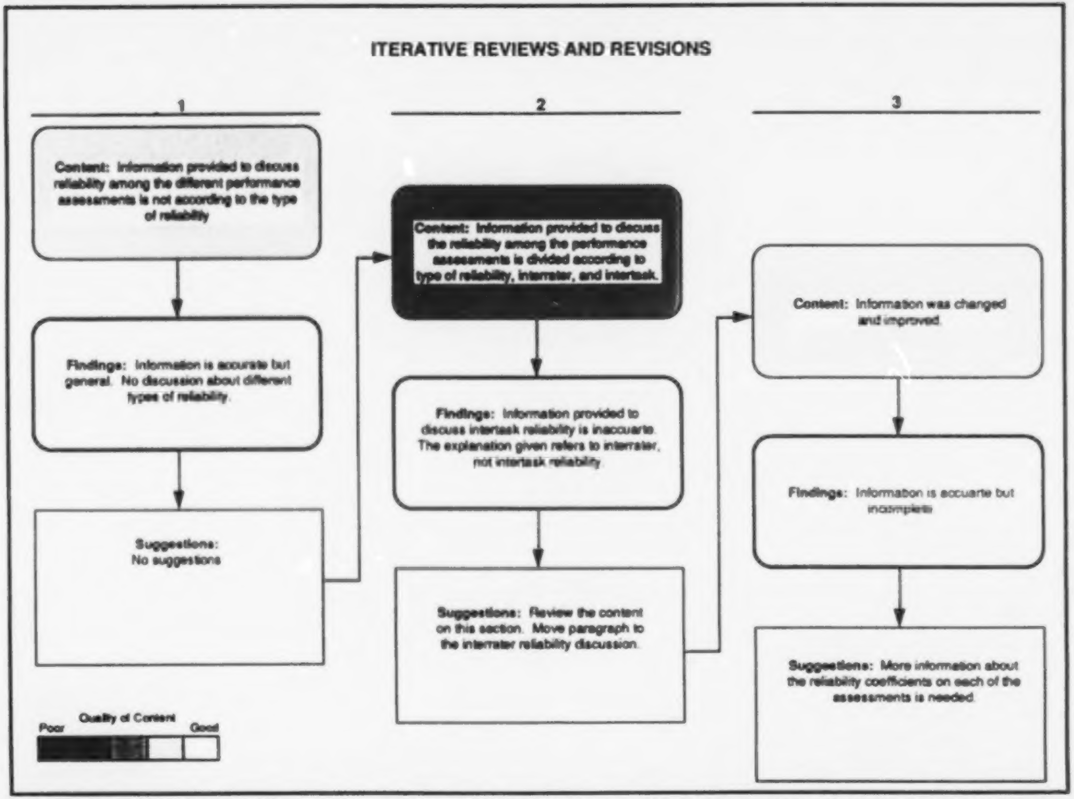
Good

☐ Information on the topic was accurate and complete. It included an adequate instructional plan for delivering the topic.

Figure 6. Criteria used to evaluate the quality of the content and shade the boxes in the flowcharts.

Findings for the same topic, "Intertask Reliability of Performance Assessments," are presented in Figure 7. This topic is discussed in different parts of the content through different activities (see Appendix C). One of the relevant activities is the "Selection Exercise". Here participants have the opportunity to review and apply knowledge acquired during the program by selecting among four different performance assessments on electricity.

Based on the reviews, done by the evaluator and the Coordinator of the Level I training, the evaluation pointed out that the content presented to facilitators for discussing intertask reliability in the selection exercise was inaccurate. The impact of the evaluation findings was immediate and new and



Content: Information provided to discuss reliability among the different performance assessments is not according to the type of reliability

Findings: Information is accurate but general. No discussion about different types of reliability.

Suggestions: No suggestions

Content: Information provided to discuss the reliability among the performance assessments is divided according to type of reliability, interrater, and intertask.

Findings: Information provided to discuss intertask reliability is inaccuarte. The explanation given refers to interrater, not intertask reliability.

Suggestions: Review the content on this section. Move paragraph to the interrater reliability discussion.

Content: Information was changed and improved.

Findings: Information is accuarte but incomplete.

Suggestions: More information about the reliability coefficients on each of the assessments is needed.

Figure 7. Quality of the content of the topic "Intertask Reliability" provided in the activity "Selection of Performance Assessments."

BEST COPY AVAILABLE

accurate information was included. It is important to mention that the iterative reviews and revisions to the program material did not correspond to each iterative tryout. Many reviews and revisions were carried out before tryout 2 as a consequence of the modifications made to the program.

Future Dissemination Alternatives Based on the Evaluation of the Prototype TEP

Although the findings are encouraging, additional work is needed. For example, the length of the program has been a major concern. The difficulty of scheduling three-day workshops with teachers during the school year called for a variety of alternative schedules. Outside tryout 3 showed that the program could be delivered in two full days. The project has also developed two optional programs based on the three-day prototype program: A one-day workshop, and a three-hour workshop. Both have already been successfully implemented in different sites with project staff as facilitators (i.e., the principal investigators and the Coordinator of the Level I program).

A final dissemination concern focuses on the facilitators. It is important to remember that even though the outcomes indicated a positive program effect for these tryouts, direct observation revealed that facilitators delivered inaccurate information when discussing participants' questions or in talking in small/large group discussions. In the final analysis, facilitators' knowledge should be more than just the program content. They need a solid background in hands-on science and performance assessments. For example, facilitators need to know how performance are developed and how their psychometric characteristics are tested.

Unfortunately, this pilot test did not examine the "training program" for facilitators. Clearly improvements are needed. Nevertheless, information obtained from these tryouts should help to develop a set of training guidelines. For example, it is clear that first, facilitators should experience the program as participants. Moreover, training should focus on those topics that the evaluation has revealed to be particularly difficult for facilitators. Training should also provide facilitators with the opportunity to administer, score, and interpret performance assessments so that they feel comfortable in explaining ideas and answering questions. Finally, training should also incorporate information about performance assessments from journal articles and magazines.

Because only two sites were used to tryout the program it is not possible to delimit the range of settings most appropriate for implementing the program. However, information accumulated through the formative evaluation process suggests that settings in which the school district already recognizes that alternative assessment is a crucial part of hands-on science instruction are the most suitable sites in which to implement the program. This does not mean that the program cannot be implemented in settings only planning or starting to change toward a hands-on science instructional approach. It only means that the benefits from the TEP will be more obvious for those participants who are already aware of the necessity of this type of assessment in order to be congruent with new forms of instruction.

The program is also suitable for those settings in which the district is beginning a system of development and implementation of performance assessments. In these settings (e.g., resource centers), this TEP can be seen as the first step before moving to a TEP that can help them develop performance assessments.

Based on the study findings one alternative that has been considered as a way to solve the problem of "facilitator expertise" is to transfer the program to private-sector and government-funded organizations that will take responsibility for training teachers and administrators. This means that the prototype TEP will be made available to "skilled users" at research and development centers, and to a few school districts whose staff have received extensive training in its use.

Conclusions

A central task during the evaluation of a prototype program is to study how delivery and outcomes vary from site to site (e.g., Cronbach, 1982). Evaluation of prototype programs should provide information about what to expect and what to do when the TEP program becomes fully operational.

Evaluation findings across three tryouts showed that the program was successfully implemented in outside sites. In other words, it is "robust" (Berk & Rossi, 1990). It produced similar results with different facilitators despite the variations in the ways the program was implemented. However, if success is defined in terms of the accuracy of all the information delivered during the implementation of the program, the program was less successful.

Still, evaluation results encourage the dissemination of the program at different sites. Very well trained facilitators with extensive knowledge of hands-on instruction and performance assessments will help to deliver an accurate and effective TEP.

The approach used for the evaluation study proved to be helpful in achieving the evaluation goals. The iterative process provided the opportunity to test modifications made to the program on the basis of the evaluation findings and to accumulate knowledge about the program. This knowledge contributed to a better understanding of the TEP by pinpointing the conditions needed for achieving the program's goals.

BIBLIOGRAPHY

Berk, R. A. & Rossi, P. H. (1990). *Thinking about program evaluation.* Newbury Park: Sage Publications.

Chinien, C. & Hlynka, D. (1993). Formative evaluation of prototype products: From expert to connoisseur. *Educational and Training Technology International, 30*(1), 60-66.

Cronbach, L. J. (1982). *Designing evaluations of educational and social programs.* San Francisco: Jossey-Bass Publishers.

Cronbach, L. J. & Associates (1980). *Toward reform of program evaluation. Aims, methods, and institutional arrangements.* San Francisco: Jossey-Bass Publishers.

Ellis, J. D. & Goulding, P. G. (1991). *Evaluation of ENLIST Micros III: Teacher Centers for improving the use of microcomputers in science instruction.* An Interim Report to the National Science Foundation.

Ellis, J. D. & Kuerbis, P. J. (1991). A curriculum for preparing science teachers to use microcomputers. *School Science and Mathematics, 91*(6), 247-254.

Fitzsimmons, S. J., Carlson, K., Burnham, R., Heinig, S., & Stoner, D. (in preparation). *A study of NSF Teacher Enhancement Program principal investigators: 1984-1989* (Draft). Cambridge, MA: Abt Associates Inc.

Gayford, C. (1987). Biotechnology 13-18: In-service training for teachers. *Journal of Biological Education, 21*(4), 281-287.

Guskey, T. & Sparks, D. (1991). What to consider when evaluating staff development. *Educational Leadership, 49*(3), 73-76.

Hurd, P. H. (1986). Perspectives for the reform of science education. *Phi Delta Kappan, 67*(5), 353-358.

Joyce, B. (1988). Training research and preservice teacher education: a reconsideration. *Journal of Teacher Education, 39*(5), 32-36.

Joyce, B. & Showers, B. (1980). Improving inservice training; The messages of research. *Educational Leadership, 37*(5), 379-385.

Knapp, M. S., Shields, P. M., ST. John, M., Zucker, A. A., & Sterns, M. S. (1988). *An approach to assessing initiatives in science education: Summary Report: Recommendations to the National Science Foundation.* SRI Project No. 089. Stanford, CA: SRI International.

Kulm, G. and Stuessy, C. (1991). Assessment in science and mathematics education reform. In G. Kulm and S. Malcom (Eds.). *Science assessment in the service of reform* (pp. 71-87). American Association for the Advancement of Science.

Nunnally, J. C. (1970). *Introduction to psychological measurement.* New York: McGraw-Hill.

Patton, M. Q. (1984). Data collection: Options, strategies, and cautions. In L. Rutman (Ed.). *Evaluation research methods: A basic guide* (2nd. ed.) (pp. 39-63). Beverly Hills: Sage Publications.

Raizen, S. A., Baron, J. B., Champagne, A. B., Haertel, E., Mullis, I. V. S., & Oakes, J. (1989). *Assessment in science education: The middle years.* Washington, DC: The National Center for Improving Science Education.

Ruiz-Primo, M. A. (1994). *Formative Evaluation of Teacher Enhancement Programs: An Approach and Case Study.* Unpublished doctoral dissertation, University of California, Santa Barbara.

Ruiz-Primo, M. A., Shavelson, R. J. & Baxter, G. P. (1993). *An Approach to Formative Evaluation for Teacher Enhancement Programs.* Paper presented at the American Educational Research Association Annual Meeting. Atlanta, Georgia, 1993.

Scriven, M. (1967). *Goals of evaluation versus roles of evaluation: Formative and summative evaluation.* In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.) Perspectives of curriculum evaluation. AERA Monographs on Curriculum Evaluation (No. 1). Chicago: Rand McNally.

Scriven, M. (1991a). *Evaluation Thesaurus* (4th ed.), Newbury Park, CA: Sage Publications.

Scriven, M. (1991b). Beyond formative and summative evaluation. In M. McLaughlin & D. C. Phillips (Eds.). *Evaluation and Education: At quarter century* (pp 19-64). Ninetieth Yearbook of National Society for the Study of Education. Chicago, IL: The University of Chicago Press.

Shadish, W. R. Jr., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation. Theories of practice.* Newbury Park: Sage Publications.

Shavelson, R. J. & Baxter, G. P. (1990). Transfer new assessment technologies to teachers and other educators. National Science Foundation Grant No. TEP 90-55443.

Shavelson, R. J., Carey, N. B., and Webb, N. M. (1990). Indicators of science achievement: Options for a powerful policy instrument. *Phi Delta Kappan, 71*(9), 692-697.

Shavelson, R. J., Copeland, W., Baxter, G. P., Decker, D. L, & Ruiz-Primo, M. A. (1994). Inservice education models for enhancing the teaching of science. In S. Fitzsimmons and L. C. Kerpelman (Eds.) *Teacher enhancement for elementary and secondary science and mathematics: Status, issues, and problems.* Washington, D.C.: National Science Foundation. Division of Research, Evaluation and Dissemination. Directorate for Education and Human Resources.

Sparks, D. & Loucks-Horsley, S. (1990). Models of staff development. In W. R. Houston, M. Haberman, & J. Sikula (Eds.) *Handbook of research on teacher education* (pp. 235-250). New York: Macmilan Publishing Company.

Tamir, P. & Amir, R. (1981). Retrospective curriculum evaluation: An approach to the evaluation of long term effects. *Curriculum Inquiry, 11*(3), 259-278.

Weiss, C. H. (1972). *Evaluation research.* Englewood Cliffs, NJ: Prentice-Hall.

Young, D. B. & Tamir, P. (1977). Finding out what students know. *The Science Teacher, 44*, 27-28.

APPENDIX A
Topics and Issues Addressed in the Program by Goal

UNDERSTANDING

Current Issues in Curriculum and Assessments

1. Current assessments practices
   - Purposes of assessments
2. Current curriculum and assessment reform
   - Characteristics of hands-on instruction
   - Hands-on instruction and performance assessments as the two sides of the same coin

Statewide Assessment Program

3. California Assessment Program--An example of a large-scale assessment program
   - CAP assumptions
   - CAP information sources

Characteristics of Performance Assessments

4. Characterization of PA
   - Definition of assessment and performance assessments
   - Differences between traditional multiple-choice tests and performance assessments
   - Variability of performance
5. Different types of tasks
   - Variety of experiences with performance assessments
   - Comparative investigations. Example used: Paper Towels Investigation
   - Component identification Example used: Mystery Powders Investigation
   - Classification tasks Example used: Leaves Task
6. Different types of scoring forms
   - Analytic scoring forms
     - Procedure-based Example used: Paper Towels scoring form
     - Evidence-based Example used: Mystery Powders scoring form
   - Holistic scoring forms
     - Rubric Example used: Leaves rubric

USE

Administration of Performance Assessments

7. Management techniques to administer performance assessments in the classroom
   - Organization of materials
   - Help from students
   - Observation during assessments
8. Embedded and End-of-Unit assessments Example: Mystery Powders embedded assessment # 2.

Scoring Performance Assessments

9. Range of alternative assessments
   - Array of performance assessments
   - Direct observation of students' performance
   - Notebooks as surrogates of direct observation
     - exchangeability
10. Practicing scoring students' performance
    - Scoring from direct observation. Example: Scoring Paper Towels (videos)
    - Scoring students' notebooks Example: Paper Towels, Mystery Powders, and Leaves

Interpreting Performance Assessments

11. Interpreting performance assessments scores
    - Summary of scores
    - Patterns of scores
    - Interpretation of patterns of scores

SELECTION

Reliability of Performance Assessments

12. Interrater reliability of performance assessments
    - High and low reliability coefficients
    - Research findings about interrater reliability
13. Intertask reliability of performance assessments
    - Research findings about intertask reliability

Validity of Performance Assessments

14. Criterion validity
    - Sensibility of performance assessments to different types of curriculum
      - Research findings about curriculum sensitivity
    - Comparison between multiple-choice traditional tests and performance assessments
      - Research findings about discriminant validity
    - Comparison between performance assessments and aptitude tests and multiple-choice and aptitude tests
      - Research findings about these correlations
15. Content validity
    - Overlapping of performance assessments and teaching unit
    - Logical judgment as evidence for content validity

Utility of Performance Assessments

16. Usefulness of scores
    - Utility of scores to evaluate students' performance
    - Utility of scores for curriculum monitoring
17. Easiness of the assessment in the classroom and the scoring form
18. Cost-effectiveness of the assessment
    - Time and cost

APPENDIX B

Sequence and Organization of the Program Content of the Prototype Program Across Tryouts 1, 2 and 3

TRYOUT 1	TRYOUT 2-3

1. **Introductions and Overview**
• Initial introductions
• Expectations and Pretest
• Workshop goals: Characteristics, use, and selection of performance assessments
• Review of agenda
	1. **Introductions and Overview**
	• Workshop opening
	• Workshop goals
	• Participants' expectations
	• Facilitators' expectations
	• Current curriculum reform in science education
	  * pre- and post-reform
	• Current assessment reform in science education
	  * pre-and post-reform
	• Workshop format

2. **Hands-on Experience with Traditional and Performance Assessments**
• Working with assessments
  * Multiple-Choice Tests Discussion of the tasks
  * Performance Assessment Circus
• Current curriculum reform
  * characteristics of hands-on science instruction
• Needed reform in assessment
  * two sides of same coin
	2. **Experience with Assessments**
	• Introduction to assessments
	• Multiple-choice tests
	• Variety of performance assessments
	• Characteristics of performance assessments
	  * performance task
	  * performance product
	  * scoring system
	• Management techniques in the classroom

3. **Assessment and Performance Assessment**
• Current assessment practices
  * four purposes of assessment
  * definition of assessment
  * definition of performance assessment
  * range of performance assessments
    - direct observation
    - notebooks
    - computer simulation
    - short-answer questions
    - new multiple-choice items
	3. **Assessments in Science Education**
	• Introduction to varieties of assessments
	  * Mini-Lecture: Variety of assessments
	    - portfolios
	    - performance assessments methods
	      - direct observation
	      - assessment notebooks
	      - computer simulation
	      - new paper-and-pencil tests
	      - exchangeability
	• Content check

4. **Hands-on Experience with a Performance Assessment**
• Observation and scoring without scoring system
• Discussion of results
  * variability in performance
  * variability in scoring
	4. **Paper Towels Investigation**
	• Observation and scoring of Paper Towels Investigation
	• Score summary form
	• Small group discussion
	• Large group discussion
	• Scoring performance
	  * Mini-Lecture: Scoring performance
	    - variability in performance
	    - variability of observers' scores
	    - interrater reliability
	• Content check
	• Video of students conducting paper towels investigation

APPENDIX B (Continued)

TRYOUT 1	TRYOUT 2-3
5. Introduction to Scoring Systems • Introduction to a procedure-based scoring system for Paper Towels • Presentation of procedure-based scoring of Paper Towels * description and judgment • Practice scoring from a student's notebook	**5. Introduction to Scoring Systems** • Scoring direct observation of the Paper Towels task * Mini-Lecture: Scoring form for Paper Towels - scoring from direct observation - method for getting the towel wet - saturation - determining result - care in saturation and/or measuring - correct result - assigning a grade • Scoring the Paper Towels assessment notebook * Mini-Lecture: Assessment notebooks and scoring performance - assessment notebook - Miguel's notebook - method for getting the towel wet - saturation - determine result - care in saturation and/or measurement - correct result - assigning grade • Discussion on the scoring of assessment notebooks • Scoring Cecilia's assessment notebooks • Exchangeability of notebooks as surrogates of direct observation * Mini-Lecture: Exchangeability - notebooks as surrogates of direct observation - exchangeability - exchangeability across assessment methods
6. Technical Quality of Performance Assessments • Judging performance assessments • Criteria * reliability - interrater reliability research findings - intertask reliability research findings * validity - content validity research findings - curriculum sensitivity research findings - discriminant validity research findings * utility • Example in technical characteristics of performance assessments	**6. Mystery Powders Investigation** • Description of the Mystery Powders Unit * Mini-Lecture: Mystery Powders Unit - general description of Mystery Powders Unit - assessments in the Mystery Powders Unit • Performing the Mystery Powders embedded assessment task • Discussion of the Mystery Powders embedded assessment task * Mini-Lecture: Content validity - content validity • A scoring system for Mystery Powders Assessment Notebooks • Scoring students' mystery powders assessment notebook * Mini-Lecture: Steps in scoring - what's inside the bag - observing tests - quality of evidence score - determining total scores. - additional issues

BEST COPY AVAILABLE

APPENDIX B (Continued)

TRYOUT 1	TRYOUT 2-3
	• Additional practice scoring with Sam's assessment notebook • Introduction to interpreting performance assessments • Step 1: Compile a summary of individual scores • Step 2: Examine the summary form • The Placement of assessment in teaching units * Mini-Lecture: Assessments in the Mystery Powders Unit - general review of mystery powders assessment - embedded assessment - end-of-unit assessment
7. Hands-On Experience with Embedded Performance Assessment • Description of the Mystery Powders Unit • Performing the mystery powders embedded assessment task • Discussing the task • Scoring students' mystery powders notebooks * explanation of mystery powders scoring form • Interpreting performance assessments • Embedded and end-of-unit assessments • Genres of performance assessments * comparison * decomposition * taxonomy * description	**7. Important Qualities of Performance Assessments** • Qualities of performance assessments • Exploring qualities of performance assessments • Application to the Mystery Powders assessment • Reliability as a quality of performance assessments * Mini-Lecture: Reliability - definition - interrater reliability - intertask reliability - research findings and their implications • Content Validity as a quality of performance assessments * Mini-Lecture: Content validity - definition of content validity - questions teachers might ask about content validity - research findings and their implications • Utility as a quality of performance assessment * Mini-Lecture: Utility and practicality - definition of utility - questions teachers might ask about utility - research findings on utility - definition of practicality - questions to be asked in determining practicality - research findings on practicality

APPENDIX B (Continued)

TRYOUT 1	TRYOUT 2-3
8. Hands-On Experience with Another Performance Assessment • Performing the CAP assessment task: Leaves • Discussion of this new genre of assessment • Introduction of a Holistic Scoring System: Discussion of the concept "Rubric" * development process of the rubric • Scoring notebooks using a rubric • Some insights into the future of large-scale performance assessment--The California case • A closer look at the 1992 CAP performance assessment in science	**8. Bugs Investigation** • Judging performance assessments • The bugs investigation * Mini-Lecture: Bugs investigation - description of mealworms unit - description of bugs performance task - description of bugs performance product - description of bugs scoring system - research of bugs investigation - interrater reliability - intertask reliability - content validity - exchangeability - utility - practicality
9. Selection of Performance Assessment • Criteria for judging PA * reliability review and questions * validity review and questions * utility review and questions • Selecting performance assessments • Discussion of selection exercise • Recommendations • Conclusion to the Workshop	**9. Rubric Scoring** • Current science assessment practice • The Leaves assessment • Introduction of a Rubric scoring system - definition of rubric - development of a rubric • Scoring Jacob's notebook using a rubric • Scoring Leticia's notebook using a rubric • General discussion on using a rubric • Distinctions among scoring systems * Mini-Lecture: Distinctions among scoring systems - additional distinctions among scoring systems - scoring systems
	10. Selection of Performance Assessments • Management of performance assessments • Criteria for judging performance assessments or "Afternoon at the Improv" • Selecting performance assessments • Discussion of selection activity
	11. Review and Closing • Creation of visual representation of terms and concepts • Closing of workshop

APPENDIX C

Topics and Issues Addressed in the "New" Prototype Program by Goal

UNDERSTANDING	USE	SELECTION
Current Issues in Curriculum and Assessments 1. Current curriculum and assessment reform • Characteristics of science instruction pre- and post-science reform • Science Assessment pre- and post-science reform **Characteristics of Performance Assessments** 2. Performance Assessments Components • Performance Task • Performance Product • Scoring Form 3. Different types of tasks • Variety of experiences with performance assessments • Comparative investigations. Example used: Paper Towels Investigation • Component identification Example used: Mystery Powders Investigation • Classification tasks Example used: Leaves Task 4. Different types of scoring forms • Analytic scoring forms - Procedure-based Example used: Paper Towels scoring form - Evidence-based Example used: Mystery Powders scoring form • Holistic scoring forms - Rubric Example used: Leaves rubric 5. Performance Assessments Methods • Direct Observation • Notebooks • Computer Simulation • New Paper-and-Pencil Tests 6. Embedded and End-of-Unit Performance Assessments • Example: Mystery Powders Embedded Assessment #2	**Administration of Performance Assessments** 7. Management techniques to administer performance assessments in the classroom **Scoring Performance Assessments** 8. Practicing scoring students' performance • Scoring from direct observation. Example: Scoring Paper Towels (videos) • Scoring students' notebooks Example: Paper Towels, Mystery Powders, and Leaves **Interpreting Performance Assessments** 9. Strategy to interpret Performance Assessments Scores • Summary of scores • Patterns of scores • Interpretation of patterns of scores Example: Interpreting Mystery Powders #2	**Reliability of Performance Assessments** 10. Interrater reliability of performance assessments • High and low reliability coefficients • Research findings about interrater reliability 11. Intertask reliability of performance assessments • Research findings about intertask reliability **Validity of Performance Assessments** 12. Content validity • Overlapping of performance assessments and teaching unit • Logical judgment as evidence for content validity **Exchangeability of Performance Assessments** 13. Exchangeability of: • Notebooks to direct observation • Computer simulation to direct observation • New paper-and-pencil to direct observation **Utility of Performance Assessments** 14. Usefulness of scores • Utility of scores to evaluate students' performance • Utility of scores for curriculum monitoring **Practicality of Performance Assessments** 15. Easiness of the assessment in the classroom and the scoring form 16. Cost-effectiveness of the assessment • Time and cost

BEST COPY AVAILABLE

Recent publications:

139	Erkki Pehkonen – Lea Lepmann 1995. Vergleich der Auffassung von Lehrern über den Mathematikunterricht in Estland und Finnland.	50,-
140	Pertti Kansanen 1995. (Hg.) Diskussionen über einige pädagogische Fragen V.	50,-
141	Erkki Pehkonen 1995. (ed.) NORMA-94 Conference Proceedings of the Nordic Conference on Mathematics Teaching (NORMA-94) in Lahti 1994.	50,-
142	Marjatta Takala 1995. "They say I'm stupid, but I just don't HEAR". Hearing-impaired Adults' view of Finnish Society.	50,-
143	Leena Kaukinen 1995. Elongation Behaviour of Elastic Stitch Types in Household Sewing Machines. Stretch-Stitches versus Serger Overlock Stitches.	50,-
144	Hannu Kuitunen – Veijo Meisalo 1995. Luovan ongelmanratkaisun leviäminen koulun työtapana. Tapaustutkimus LOTTO-projektin toiminnasta.	50,-
145	Pertti Kansanen (Ed.) 1995. Discussions on Some Educational Issues VI.	50,-
146	Jari Metsämuuronen 1995. Harrastukset ja omaehtoinen oppiminen: Sitoutuminen, motivaatio ja coping. Teoreettinen tausta, rakenneanalyysi ja sitoutuminen.	50,-
147	Heikki Kynäslahti – Anneli Lauriala – Markku Suortamo (toim.) 1995. Kokeileva POKO. Luokanopettajan poikkeuskoulutus opettajankoulutuksen kehittäjänä.	50,-
148	Anna-Liisa Sysiharju 1995. Naisasian tytär – muuttuvien kotien tuki 1891 – 1990. Vuosisata kotitalousopettajien koulutusta Helsingissä.	100,-
149	Jussi T. Koski 1995. Horisonttiensulautumisia: Keskustelua Hans-Georg Gadamerin kanssa hermeneutiikasta, kasvamisesta, tietämisestä ja kasvatustieteestä.	50,-
150	Seppo Tella (toim.) 1995. Juuret ja arvot: Etnisyys ja eettisyys – aineen opettaminen monikulttuurisessa oppimisympäristössä. Ainedidaktiikan symposiumi Helsingissä 3.2.1995.	50,-
151	Matti Erätuuli – Christer Nylèn 1995. The improvement of school leadership: Co-operation between Russian, Swedish and Finnish principals. Part 1: The background, context and the principals' job descriptions.	50,-
152	Erkki Pehkonen 1995. Pupils' View of Mathematics: Initial report for an international comparison project.	50,-
153	Kirsi Tirri 1995. Ala-asteen oppilaiden identifioimien moraaliongelmien teemat ja sosiaaliset suhteet.	50,-
154	Pirjo Jauhiainen 1995. Opetussuunnitelmatyö koulussa: Muuttuuko yläasteen opettajan työ ja ammattikuva?	50,-

BEST COPY AVAILABLE

UNIVERSITY OF HELSINKI
DEPARTMENT OF TEACHER EDUCATION

ISBN 951-45-7034-0
ISSN 0359-4203
Yliopistopaino
1995